

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 23:33:55 ; Search time 113.83 Seconds

(without alignments)
16.268 Million cell updates/sec

Title: US-09-164-862B-1

Perfect score: 139

Sequence: 1 VKLVCTYSWQYREGDGSXRPDAL 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_1101.*

1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
| 1 | 137 | 98.6 | 25 | 16 | AA1980 |
| 2 | 137 | 98.6 | 25 | 16 | AA1981 |
| 3 | 137 | 98.6 | 25 | 21 | AA1991 |
| 4 | 136 | 97.8 | 43 | 18 | AA1992 |
| 5 | 136 | 97.8 | 126 | 21 | AA1993 |
| 6 | 136 | 97.8 | 321 | 9 | AA1994 |
| 7 | 134 | 96.4 | 383 | 21 | AA1995 |
| 8 | 133 | 95.7 | 383 | 18 | AA1996 |
| 9 | 95 | 68.3 | 16 | 18 | AA1997 |
| 10 | 86 | 61.9 | 16 | 18 | AA1998 |
| 11 | 84 | 60.4 | 455 | 22 | AA1999 |

| | | | | | | |
|----|------|------|-----|----|--------|--------------------|
| 12 | 84 | 60.4 | 476 | 22 | AA1980 | Disease treatment |
| 13 | 83 | 59.7 | 373 | 19 | AA1981 | Human chitinase pr |
| 14 | 83 | 59.7 | 373 | 19 | AA1982 | Human chitinase pr |
| 15 | 83 | 59.7 | 373 | 20 | AA1983 | Clone of the C-ter |
| 16 | 83 | 59.7 | 373 | 20 | AA1984 | Chitinase amino ac |
| 17 | 83 | 59.7 | 387 | 18 | AA1985 | Human 39 kDa chiti |
| 18 | 83 | 59.7 | 464 | 18 | AA1986 | Human chitinase |
| 19 | 83 | 59.7 | 466 | 18 | AA1987 | Human chitinase |
| 20 | 83 | 59.7 | 466 | 18 | AA1988 | Human chitinase |
| 21 | 83 | 59.7 | 466 | 19 | AA1989 | Human chitinase pr |
| 22 | 83 | 59.7 | 466 | 19 | AA1990 | Human chitinase pr |
| 23 | 83 | 59.7 | 466 | 20 | AA1991 | MO-218 clone of hu |
| 24 | 83 | 59.7 | 466 | 20 | AA1992 | MO-13B clone of hu |
| 25 | 83 | 59.7 | 466 | 22 | AA1993 | Human chitinase pr |
| 26 | 83 | 59.7 | 466 | 22 | AA1994 | Human chitinase pr |
| 27 | 78 | 56.1 | 89 | 22 | AA1995 | Human ESR encoded |
| 28 | 78 | 56.1 | 108 | 21 | AA1996 | Human secreted pro |
| 29 | 78 | 56.1 | 385 | 19 | AA1997 | Amino acid sequenc |
| 30 | 78 | 56.1 | 385 | 19 | AA1998 | Human cartilage gp |
| 31 | 78 | 56.1 | 416 | 19 | AA1999 | Human cartilage gp |
| 32 | 78 | 56.1 | 421 | 19 | AA2000 | Amino acid sequenc |
| 33 | 78 | 56.1 | 421 | 19 | AA2001 | Amino acid sequenc |
| 34 | 78 | 56.1 | 423 | 18 | AA2002 | Human cartilage gl |
| 35 | 77 | 55.4 | 68 | 21 | AA2003 | Human secreted pro |
| 36 | 75 | 54.0 | 718 | 16 | AA2004 | Murine oviduct spe |
| 37 | 70 | 50.4 | 377 | 22 | AA2005 | Disease treatment |
| 38 | 70 | 50.4 | 398 | 22 | AA2006 | Disease treatment |
| 39 | 69 | 49.6 | 25 | 21 | AA2007 | House dust mite al |
| 40 | 66 | 47.5 | 537 | 16 | AA2008 | Bovine oviduct spe |
| 41 | 65 | 46.8 | 668 | 16 | AA2009 | Hamster oviduct sp |
| 42 | 64 | 46.0 | 16 | 18 | AA2010 | Human cartilage gl |
| 43 | 59 | 42.4 | 554 | 18 | AA2011 | Manduca sexta larv |
| 44 | 59 | 42.4 | 554 | 21 | AA2012 | Manduca sexta gut |
| 45 | 57.5 | 41.4 | 440 | 21 | AA2013 | Drosophila melanog |

ALIGNMENTS

RESULT 1

AA1980

ID AA1980 standard; peptide; 25 AA.

XX AA1980

XX AA1980

XX 27-SEP-1995 (first entry)

XX YKL-40 N-terminal peptide.

XX YKL-40 N-terminal peptide.

XX YKL-40; N-terminal peptide; connective tissue degradation;

XX osteoporosis; tumour metastasis; polysaccharide hydrolase;

XX chitinase; joint disease; competitive immunoassay.

XX Homo sapiens.

XX Key Location/Qualifiers

XX Misc-difference 20

XX /note= "undefined"

XX WO9502188-A.

XX 19-JAN-1995.

XX 12-JUL-1993; 93WO-US06579.

XX 12-JUL-1993; 93WO-US06579.

XX (REGC) UNIV CALIFORNIA.

XX Johansen JS, Price PA;

XX WPI; 1995-066992/09.

PT Assay for YKL-40 as a marker for the degradation of connective
 PT tissue - is a competitive immunoassay, and can be used to
 PT identify joint disease and the extent of tumour metastasis
 XX
 PS Disclosure; Page 33; 51pp; English.
 XX
 CC AAR70751, AAR70752 and AAR70753 are the N-terminal, and internal
 CC peptides A and B respectively of the human osteosarcoma cell
 CC line MG63 derived protein YKL-40. These peptides show a sequence
 CC homology with a bacterial polysaccharide hydrolase (chitinase),
 CC this suggests that the protein YKL-40 is involved in connective
 CC tissue degradation (CTD). By performing a competitive immunoassay
 CC for YKL-40 (as a marker of CTD) joint disease, osteoporosis and
 CC the extent of tumour metastasis can be diagnosed.
 XX
 XX

SQ Sequence 25 AA;

Query Match 98.6%; Score 137; DB 16; Length 25;
 Best Local Similarity 100.0%; Pred. No. 8.1e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKLVCYITWSQYREGDGSXFPDAL 25
 DB 1 yklvcyitwsqyregdgsxfpdal 25
 |||||

RESULT 2
 AAR70745
 ID AAR70745 standard; Peptide; 25 AA.
 AC AAR70745;

DT 12-AUG-1995 (first entry)
 DE YKL-40 N-terminal sequence.

XX YKL-40; diagnosis; prognosis; therapy; breast cancer; metastasis;
 KW marker; joint disease; connective tissue.
 XX Homo sapiens.

OS
 FH Key Location/Qualifiers
 FT Misc-difference 20 /note= "not specified"
 FT

PN WO9501995-A.

XX 19-JAN-1995.

XX 08-JUL-1994; 94WO-US07754.

XX 09-JUL-1993; 93US-0089989.

XX (REGC) UNIV CALIFORNIA.

XX Johansen JS, Price PA;

XX WPI; 1995-066866/09.

XX Use of YKL-40 and anti-YKL-40 antibodies - for developing prods.
 PT for diagnosis prognosis and therapy of diseases involving
 PT connective tissue degradation.
 XX

PS Disclosure; Page 64; 88pp; English.

XX YKL-40 (40 kDa) was purified from human osteosarcoma MG63 cells.
 CC The N-terminal sequence is shown in AAR70745; the full coding region
 CC of the YKL-40 gene is given in AAR70745. Homology of the N-terminal
 CC and 2 internal peptides (AAR70746-47) with a bacterial polysaccharide
 CC hydrolase suggests that YKL-40 degrades polysaccharide components
 CC of connective tissue. YKL-40 is a marker of e.g. metastatic breast
 CC cancer and inflammatory or degenerative joint diseases.
 CC

XX SQ Sequence 25 AA;
 Query Match 98.6%; Score 137; DB 16; Length 25;
 Best Local Similarity 100.0%; Pred. No. 8.1e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 YKLVCYITWSQYREGDGSXFPDAL 25
 DB 1 yklvcyitwsqyregdgsxfpdal 25
 |||||

RESULT 3

AAI79491
 ID AAI79491 standard; Peptide; 25 AA.

XX AAI79491;

DT 01-AUG-2000 (first entry)

XX Human cancer marker YKL-40 N-terminal peptide.
 DE
 XX

KW YKL-40; human; lung cancer; bronchus cancer; colorectal cancer;
 KW prostate cancer; breast cancer; pancreas cancer; stomach cancer;
 KW ovary cancer; bladder cancer; brain cancer; oesophagus cancer;
 KW cervix cancer; melanoma; uterine endometrial cancer;
 KW oral cavity cancer; pharynx cancer; liver cancer; kidney cancer;
 KW biliary tract cancer; small bowel cancer; appendix cancer;
 KW salivary gland cancer; thyroid gland cancer; testis cancer;
 KW adrenal gland cancer; osteosarcoma; chondrosarcoma; liposarcoma;
 KW malignant fibrous histiocytoma; infection; pneumonia; meningitis;
 KW arteritis; rheumatoid arthritis; osteoarthritis; fibrosis;
 KW liver cirrhosis; marker; diagnosis; prognosis.

XX Homo sapiens.

OS
 FH Key Location/Qualifiers
 FT Misc-difference 20 /note= "unidentified amino acid residue"

FT
 XX WO200019206-A1.

XX 06-APR-2000.

XX 29-SEP-1999; 99WO-US22615.

XX 01-OCT-1998; 98US-0164862.

XX (REGC) UNIV CALIFORNIA.

XX Price PA, Johansen JS;

XX WPI; 2000-303485/26.

XX Novel methods for detecting cancers and evaluating the prognosis of
 PT cancer using YKL-40 as a marker of cancer -
 XX
 XX Disclosure; Page 101; 111pp; English.

XX This peptide represents the N-terminal sequence of human YKL-40
 CC mature polypeptide. YKL-40 is a 40 kDa protein having Tyr, Lys and
 CC Leu as its N-terminal residues (hence, YKL-40). The protein was
 CC obtained from osteosarcoma cell line MG63. It is a mammalian
 CC member of the chitinase family, and is suggested to degrade the
 CC polysaccharide components in connective tissue and/or is a lectin
 CC that binds to specific glycan structures in the extracellular
 CC environment of cells. YKL-40 is useful as a marker for the
 CC presence or absence of a cancer and for the prognosis of a cancer.
 CC A claimed method for estimating survival length of cancer patients
 CC comprises obtaining a biological sample from the patient and
 CC measuring the level of YKL-40, a higher level than in healthy
 CC humans being indicative of reduced survival expectancy. The

CC biological sample is obtained from a cancer patient having at least
 CC a preliminary diagnosis of cancer selected from lung, bronchus,
 CC colorectal, prostate, breast, pancreas, stomach, ovary, urinary
 CC bladder, brain, central nervous system, peripheral nervous system,
 CC oesophagus, cervix, melanoma, uterine endometrial, oral cavity,
 CC pharynx, liver, kidney, biliary tract, small bowel, appendix,
 CC salivary gland, thyroid gland, testes, or adrenal gland cancer, or
 CC osteosarcoma, chondrosarcoma, liposarcoma, or malignant fibrous
 CC histiocytoma. Levels of the YKL-40 marker are elevated in
 CC pathologies associated with tissue remodeling, e.g. degenerative
 CC bone diseases such as rheumatoid arthritis, osteoarthritis, fibrosis,
 CC cirrhosis of the liver, and cancer, especially breast, colon,
 CC prostate, or lung cancer. The marker can be used to identify high
 CC risk patients, and so allow selection of appropriate therapeutic
 CC regimens. The methods may also be used to detect bacterial
 CC infections, such as bacterial pneumonia and meningitis, as these
 CC cause an elevation in YKL-40 levels, as well as diseases
 CC characterized by macrophage activation, e.g. giant cell arteritis.
 CC The YKL-40 marker may also be used to evaluate treatment efficacy,
 CC to check for recurrence of a cancer, to monitor terminal phase
 CC patients, and to check the efficacy of surgical removal of a
 CC primary tumor. The methods allow estimation of the survival time
 CC of patients with cancers, especially prostate, lung or colorectal
 CC cancer, where the colorectal cancer is Duke's stage A, B, C, or D.
 XX
 SQ Sequence 25 AA;

Query Match 98.6%; Score 137; DB 21; Length 25;
 Best Local Similarity 100.0%; Pred. No. 8.1e-13;
 Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 YKLVCYTSMOYREGDGSXFPDAL 25
 |||||
 Db 1 YKLVCYTSMOYREGDGSXFPDAL 25

RESULT 4

ID AAW26750 standard; Protein; 43 AA.

AC AAW26750;

DT 11-MAY-1998 (first entry)

DE Human cartilage glycoprotein 39 (HC gp-39) autoantigen.

KW Human cartilage glycoprotein 39; HC gp-39; autoantigen;
 KW autoimmune disease; rheumatoid arthritis; inflammation;
 KW arthritogenic protein; immunotherapy; therapy.

OS Homo sapiens.

PN WO9740149-A1.

PD 30-OCT-1997.

PF 15-APR-1997; 97WO-EP01903.

PR 18-APR-1996; 96US-0634493.

PA (ALKU) AKZO NOBEL NV.

PI Boots AMH, Bos ES, Verheijden GPM;

DR WPI; 1997-535833/49.

XX Autoantigen proteins homologous to human cartilage glycoprotein 39 -
 PT induce arthritis and provide antigen-specific treatment of articular
 PT cartilage destruction in autoimmune diseases e.g. rheumatoid
 PT arthritis
 XX
 PS Claim 1; Page 27; 35pp; English.

XX This protein comprises human cartilage glycoprotein 39 (HC gp-39)
 CC autoantigen, a target autoantigen in rheumatoid arthritis patients,
 CC which activates specific T cells, thus causing or mediating the
 CC inflammatory process. HC gp-39 and proteins having an amino acid
 CC sequence which exhibits at least 50% homology with HC gp-39, can be
 CC used in the antigen-specific treatment of articular cartilage
 CC destruction in autoimmune diseases in mammals to induce systemic
 CC tolerance of the immune system. The autoantigen HC gp-39, and
 CC arthritogenic proteins, e.g. bovine whey protein (see AAW26751) or
 CC human YKL-39 protein, having sequences that exhibit at least 50%
 CC homology with HC gp-39, are also suitable to induce arthritis in
 CC animals, preferably mice, e.g. for use in drug screening. The
 CC invention also relates to pharmaceutical compositions comprising
 CC the autoantigen and/or arthritogenic proteins, and a diagnostic
 CC method for the detection of autoreactive T cells in a test sample.
 CC HC gp-39 has a specific tolerising effect on the autoreactive T cells
 CC and does not cause toxic side-effects.
 XX
 SQ Sequence 43 AA;

Query Match 97.8%; Score 136; DB 18; Length 43;
 Best Local Similarity 96.0%; Pred. No. 2e-12;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 YKLVCYTSMOYREGDGSXFPDAL 25
 |||||
 Db 1 YKLVCYTSMOYREGDGSXFPDAL 25

RESULT 5

ID AAG00227 standard; Protein; 126 AA.

AC AAG00227;

DT 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 4308.

KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.

OS Homo sapiens.

PN EP1033401-A2.

PD 06-SEP-2000.

PF 21-FEB-2000; 2000EP-0200610.

PR 26-FEB-1999; 99US-0122487.

PA (GEST) GENSET.

PI Dumas Milne Edwards J, Duclert A, Giordano J;

DR WPI; 2000-500381/45.

DR N-PSDB; AAC00233.

PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX
 PS Claim 13; SEQ ID 4308; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for

CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream
 CC regulatory sequences and to design expression and secretion vectors.
 XX
 SQ Sequence 126 AA;

Query Match 97.8%; Score 136; DB 21; Length 126;
 Best Local Similarity 96.0%; Pred. No. 6.2e-12;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKLYCYTWSQYREGDGSXFPDAL 25
 DB 22 YKLYCYTWSQYREGDGSXFPDAL 46

RESULT 6
 AAP81342
 ID AAP81342 standard; protein; 321 AA.
 AC
 AC
 XX
 XX
 19-OCT-1990 (first entry)
 XX
 DE Polypeptide involved in protective mechanisms.
 XX
 KW Immune response; cell growth.
 XX
 PN JP63032898-A.
 XX
 PD 30-JAN-1988.
 XX
 PF 16-JUL-1986; 86JP-0167518.
 XX
 PR 16-JUL-1986; 86JP-0167518.
 XX
 XX (DAIN) DAINIPPON PHARM KK.
 PA
 XX WPI; 1988-068419/10.
 DR P-PSDB; AAP81342.
 XX
 XX New polypeptide and DNA encoding it - related to protective
 PT mechanisms such as immune response etc.
 PT
 PS Disclosure; ; p; Japanese.
 XX
 CC This polypeptide is involved in protective mechanisms such as immune
 CC response, cell growth and activation of protective functions.
 CC
 XX
 SQ Sequence 321 AA;

Query Match 97.8%; Score 136; DB 9; Length 321;
 Best Local Similarity 96.0%; Pred. No. 1.6e-11;
 Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKLYCYTWSQYREGDGSXFPDAL 25
 DB 21 YKLYCYTWSQYREGDGSXFPDAL 45

RESULT 7
 AAB03442
 ID AAB03442 standard; Protein; 383 AA.
 XX
 AC
 XX
 XX
 DT 03-JAN-2001 (first entry)
 XX

DE Gp38k protein sequence.
 XX
 KW Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
 KW cancer; vascular trauma; vascular disease; atherosclerosis; restenosis;
 XX
 OS Unidentified.
 XX

Key Location/Qualifiers
 FH Peptide 1..24
 FT /label= signal_peptide
 FT 30..37
 FT Region
 FT /note= "peptide antibody"
 FT Modified-site 59..61
 FT /label= glycosylation_site
 FT Binding-site 68..75
 FT /label= leucine_zipper
 FT Binding-site 82..89
 FT /label= leucine_zipper
 FT Active-site 131..136
 FT /label= chitinase_active_site
 FT Binding-site 143..146
 FT /label= heparin_binding_site
 FT Binding-site 147..154
 FT /label= hyaluronic_acid_binding_site
 FT Binding-site 262..270
 FT /label= hyaluronic_acid_binding_site
 FT Binding-site 278..281
 FT /label= glycosaminoglycan_binding_site
 FT Region 354..357
 FT /note= "acidic region"
 FT Binding-site 368..376
 FT /label= hyaluronic_acid_binding_site
 XX
 PN WO200034469-A1.
 XX
 PD 15-JUN-2000.
 XX
 PF 10-DEC-1999; 99WO-US29262.
 XX
 PR 11-DEC-1998; 98US-0111856.
 XX
 XX (UUNY) UNIV NEW YORK STATE RES FOUND.
 PA
 XX
 PI Millis AJT;
 XX
 DR WPI; 2000-431300/37.
 XX
 XX Clusterin and gp38k-related peptide capable of altering cell migration
 PT useful for treating atherosclerosis, cancer and stenosis following
 PT vascular trauma or disease
 XX
 PS Disclosure; Fig 2; 43pp; English.
 XX

The present sequence is the protein sequence of gp38k. Gp38k, a
 CC chemoattractant, is essential for the migration of vascular smooth muscle
 CC cells (VSMC). The gene and protein can, therefore, be used to promote
 CC wound healing, angiogenesis and vasculogenesis, in the treatment of
 CC stenosis following vascular trauma or disease and to treat
 CC atherosclerosis, and antisense sequences can be used to treat cancer, as
 CC angiogenesis is vital for tumour survival.
 XX
 SQ Sequence 383 AA;

Query Match 96.4%; Score 134; DB 21; Length 383;
 Best Local Similarity 92.0%; Pred. No. 3.8e-11;
 Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKLYCYTWSQYREGDGSXFPDAL 25
 DB 22 YKLYCYTWSQYREGDGSXFPDAL 46

KW HC gp-39; rheumatoid arthritis; epitope.
 OS Synthetic.
 OS Homo sapiens.
 XX WO9740068-A1.
 PN 30-OCT-1997.
 PD 22-APR-1997; 97WO-EP02051.
 XX 24-APR-1996; 96EP-0201106.
 XX (ALKU) AKZO NOBEL NV.
 PA Boots AMH, Verheijden GFM;
 PI WPI; 1997-535775/49.
 XX Peptide suitable for use in antigen specific immunosuppressive
 PT therapy - resembles or mimics epitope present on HC gp-39, so
 PT inducing systemic immunological tolerance to rheumatoid arthritis
 PT auto:antigen
 XX Claim 2; Page 76; 82pp; English.
 PS The present sequence represents a peptide which resembles or mimics an
 XX epitope present on human cartilage glycoprotein 39 (HC gp-39), an
 CC autoantigen in rheumatoid arthritis. The invention relates to peptides
 CC consisting of 16-55 amino acid residues comprising at least one of the
 CC following 19 sequences: LVCYITWS; FLCTHIYS; ILYSFANIS; LKTLISVGG;
 CC FKSVPPPT; PDGLDLAWL; LYFGRDKQ; YDIKISQH; LDFISIMTV; FLSIMTYDF;
 CC FRCQENSP; FAVGYMLRL; MRLGAPAS; LAYVEICDF; LRGATVHR; YLKDRQLAG;
 CC LAGAMWAL; VWALDLDDF; or LDLDLDFQS. They can be used medically in
 CC antigen specific immunosuppressive therapy, particularly in the treatment
 CC of T-cell mediated destruction of articular cartilage in autoimmune
 CC diseases (e.g. rheumatoid arthritis). They can also be used to detect
 CC activated autoreactive T cells in an individual. The peptides have a
 CC specific effect on the autoreactive T cells, thus leaving the other
 CC components of the immune system intact, unlike the non-specific
 CC suppressive effect of immunosuppressive drugs, and do not cause toxic
 CC side effects. The peptides are predominantly recognised by autoreactive
 CC T cells from rheumatoid arthritis patients, but rarely by those from
 CC healthy donors.
 XX
 SQ Sequence 16 AA;

Query Match 61.9%; Score 86; DB 18; Length 16;
 Best Local Similarity 93.8%; Pred. No. 8.6e-06;
 Matches 15; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 7 YTSWSQYREGDGSXFP 22
 Db 1 YTSWSQYREGDGSXFP 16

RESULT 11
 AAG62541
 ID AAG62541 standard; protein; 455 AA.
 XX
 AC AAG62541;
 XX
 DT 24-AUG-2001 (first entry)
 XX
 DE Disease treatment related protein SEQ ID NO: 1.
 XX
 KW Disease treatment; infection; chronic occlusive pulmonary disease;
 KW bronchial asthma.
 XX
 OS Homo sapiens.
 XX
 PN WO200136633-A1.

XX 25-MAY-2001.
 XX 14-NOV-2000; 2000WO-JP08015.
 PF 15-NOV-1999; 99JP-0324467.
 XX (TAKE) TAKEDA CHEM IND LTD.
 PA Nakanishi A, Morita S;
 PI WPI; 2001-397791/42.
 XX New proteins, peptides and DNA for treatment of bronchial asthma,
 XX chronic occlusive lung disease and infectious disease -
 PS Claim 1; Page 97-99; 114pp; Japanese.
 XX The present invention provides the sequence of a protein which can be
 CC used in the treatment and prevention of infectious diseases. Inhibitors
 CC of the protein can be used to treat bronchial asthma and chronic
 CC occlusive pulmonary disease. The present sequence is a protein described
 CC in the exemplification of the invention.
 XX SQ Sequence 455 AA;

Query Match 60.4%; Score 84; DB 22; Length 455;
 Best Local Similarity 56.5%; Pred. No. 0.00056;
 Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 YKLVCYITWSQYREGDGSXFP 23
 Db 1 YQICYFTNWADYRPGIgrfmpd 23

RESULT 12
 AAG62543
 ID AAG62543 standard; protein; 476 AA.
 XX
 AC AAG62543;
 XX
 DT 24-AUG-2001 (first entry)
 XX
 DE Disease treatment related protein SEQ ID NO: 5.
 XX
 KW Disease treatment; infection; chronic occlusive pulmonary disease;
 KW bronchial asthma.
 XX
 OS Homo sapiens.
 XX
 PN WO200136633-A1.
 XX
 PD 25-MAY-2001.
 XX
 PF 14-NOV-2000; 2000WO-JP08015.
 XX
 PR 15-NOV-1999; 99JP-0324467.
 XX
 PA (TAKE) TAKEDA CHEM IND LTD.
 XX
 PI Nakanishi A, Morita S;
 XX
 DR WPI; 2001-397791/42.
 XX
 PT New proteins, peptides and DNA for treatment of bronchial asthma,
 PT chronic occlusive lung disease and infectious disease -
 PS Example 5; Page 101-103; 114pp; Japanese.
 XX
 CC The present invention provides the sequence of a protein which can be
 CC used in the treatment and prevention of infectious diseases. Inhibitors
 CC of the protein can be used to treat bronchial asthma and chronic

Search completed: December 7, 2001, 00:23:14
Job time: 2959 sec

SQ Sequence **373 AA;**

Query Match 59.7%; Score 83; DB 20; Length 373;
Best Local Similarity 54.2%; Pred. No. 0.00063;
Matches 13; Conservative 5; Mismatches 6;
Indels 0; Gaps 0;

FEATURE: OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM


```
RESULT 8
US-08-486-839-6
; Sequence 6, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; production, its use for decomposing chitin, its use
; in therapy or prophylaxis against infection diseases.
; TITLE OF INVENTION: 16
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-486-839-6

Query Match 59.7%; Score 83; DB 2; Length 387;
Best Local Similarity 54.2%; Pred. No. 0.00026;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 KLVCYTWSQYREGDGSXFPDAL 25
Db 23 KLVCFYTNWQYRQGEARFLPKDL 46

RESULT 9
US-09-151-011-6
; Sequence 6, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
; Production, Its Use For Decomposing Chitin, Its Use in
; Therapy or Prophylaxis Against Infection Diseases.
; TITLE OF INVENTION: 16
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-151-011-6
```

```
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morris, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-151-011-6

Query Match 59.7%; Score 83; DB 3; Length 387;
Best Local Similarity 54.2%; Pred. No. 0.00026;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 KLVCYTWSQYREGDGSXFPDAL 25
Db 23 KLVCFYTNWQYRQGEARFLPKDL 46

RESULT 10
US-09-343-623-6
; Sequence 6, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; production, its use for decomposing chitin, its use
; in therapy or prophylaxis against infection diseases.
; TITLE OF INVENTION: 16
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; SOFTWARE:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 387 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-343-623-6

Query Match          59.7%; Score 83; DB 4; Length 387;
Best Local Similarity 54.2%; Pred. No. 0.00026;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 KLVCVYTSWSQYREGDGSXFPDAL 25
Db 23 KLVCFYTNNAQYRQGEARELPKDL 46

RESULT 11
US-08-486-839-4
; Sequence 4, Application US/08486839
; Patent No. 5928928
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/486.839
; FILING DATE: 07 - June - 1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Baron, Ronald J.
; REGISTRATION NUMBER: 29,281
; REFERENCE/DOCKET NUMBER: 294-26
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-08-486-839-4

Query Match          59.7%; Score 83; DB 2; Length 466;
Best Local Similarity 54.2%; Pred. No. 0.00031;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 KLVCVYTSWSQYREGDGSXFPDAL 25
Db 23 KLVCFYTNNAQYRQGEARELPKDL 46

RESULT 12
US-09-151-011-4
; Sequence 4, Application US/09151011
; Patent No. 6057142
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A Human Chitinase, Its Recombinant
; TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
; TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron, LLP
; STREET: 6900 Jericho Turnpike
; CITY: Syosset
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11791
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/151,011
; FILING DATE: 10 - September - 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Morriss, Robert C.
; REGISTRATION NUMBER: 42,910
; REFERENCE/DOCKET NUMBER: 294-32 DIV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (516) 822-3550
; TELEFAX: (516) 822-3582
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 466 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
US-09-151-011-4

Query Match          59.7%; Score 83; DB 3; Length 466;
Best Local Similarity 54.2%; Pred. No. 0.00031;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

QY 2 KLVCVYTSWSQYREGDGSXFPDAL 25
Db 23 KLVCFYTNNAQYRQGEARELPKDL 46

RESULT 13
US-09-039-198A-2
; Sequence 2, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
```

Query Match 59.7%; Score 83; DB 4; Length 466;
Best Local Similarity 54.2%; Pred. No. 0.00031;
Matches 13; Conservative 5; Mismatches 6; Indels

Search completed: December 7, 2001, 00:24:38
Job time: 2868 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:19:41 ; Search time 401.11 Seconds
(without alignments)
17.306 Million cell updates/sec

Title: US-09-164-862B-1

Perfect score: 139

Sequence: 1 YKLVCYTSWSQYREGDSXFPDAL 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1: /cgn2_6/ptodata/1/paa/PCTUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 137 | 98.6 | 25 | US-08-089-989-1 | Sequence 1, Appli |
| 2 | 137 | 98.6 | 25 | US-09-164-862B-1 | Sequence 1, Appli |
| 3 | 137 | 98.6 | 25 | US-09-215-077A-1 | Sequence 1, Appli |
| 4 | 136 | 97.8 | 43 | US-09-158-300-10 | Sequence 10, Appli |
| 5 | 136 | 97.8 | 43 | US-09-171-562-1 | Sequence 1, Appli |
| 6 | 136 | 97.8 | 206 | US-09-760-475-2626 | Sequence 2626, Ap |
| 7 | 136 | 97.8 | 362 | PCT-US01-14827-14282 | Sequence 14282, A |
| 8 | 136 | 97.8 | 383 | US-08-755-051-3 | Sequence 3, Appli |
| 9 | 136 | 97.8 | 404 | PCT-US01-14827-14283 | Sequence 14283, A |

| | | | | | | |
|----|-----|------|-----|----|--------------------|--------------------|
| 10 | 134 | 95.4 | 383 | 18 | US-09-459-749B-17 | Sequence 17, Appli |
| 11 | 133 | 95.7 | 383 | 15 | US-09-171-562-2 | Sequence 2, Appli |
| 12 | 84 | 60.4 | 58 | 24 | US-60-195-053-1679 | Sequence 1679, Ap |
| 13 | 84 | 60.4 | 68 | 24 | US-60-177-571-4666 | Sequence 4666, Ap |
| 14 | 84 | 60.4 | 68 | 24 | US-60-177-846-3866 | Sequence 3866, Ap |
| 15 | 84 | 60.4 | 71 | 24 | US-60-171-481-1276 | Sequence 1276, Ap |
| 16 | 84 | 60.4 | 476 | 24 | US-60-242-679-1419 | Sequence 1419, Ap |
| 17 | 83 | 59.7 | 64 | 24 | US-60-163-233-3858 | Sequence 3858, Ap |
| 18 | 83 | 59.7 | 82 | 24 | US-60-162-247-3982 | Sequence 3982, Ap |
| 19 | 83 | 59.7 | 82 | 24 | US-60-189-840-7635 | Sequence 7635, Ap |
| 20 | 83 | 59.7 | 88 | 24 | US-60-196-710-7131 | Sequence 7131, Ap |
| 21 | 83 | 59.7 | 88 | 24 | US-60-196-718-7444 | Sequence 7444, Ap |
| 22 | 83 | 59.7 | 136 | 21 | US-09-760-476-1948 | Sequence 1948, Ap |
| 23 | 83 | 59.7 | 158 | 24 | US-60-207-216-836 | Sequence 836, App |
| 24 | 83 | 59.7 | 373 | 10 | US-08-663-618A-14 | Sequence 14, Appli |
| 25 | 83 | 59.7 | 373 | 10 | US-08-663-618A-15 | Sequence 15, Appli |
| 26 | 83 | 59.7 | 373 | 12 | US-08-877-599-14 | Sequence 14, Appli |
| 27 | 83 | 59.7 | 373 | 12 | US-08-877-599-15 | Sequence 15, Appli |
| 28 | 83 | 59.7 | 373 | 16 | US-09-267-574-14 | Sequence 14, Appli |
| 29 | 83 | 59.7 | 373 | 16 | US-09-267-574-15 | Sequence 15, Appli |
| 30 | 83 | 59.7 | 464 | 1 | PCT-US97-05072-2 | Sequence 2, Appli |
| 31 | 83 | 59.7 | 464 | 12 | US-08-819-791-2 | Sequence 2, Appli |
| 32 | 83 | 59.7 | 466 | 1 | PCT-US97-05072-4 | Sequence 4, Appli |
| 33 | 83 | 59.7 | 466 | 10 | US-08-663-618A-2 | Sequence 2, Appli |
| 34 | 83 | 59.7 | 466 | 10 | US-08-663-618A-4 | Sequence 4, Appli |
| 35 | 83 | 59.7 | 466 | 11 | US-08-755-051-4 | Sequence 4, Appli |
| 36 | 83 | 59.7 | 466 | 12 | US-08-819-791-4 | Sequence 4, Appli |
| 37 | 83 | 59.7 | 466 | 12 | US-08-877-599-2 | Sequence 2, Appli |
| 38 | 83 | 59.7 | 466 | 12 | US-08-877-599-4 | Sequence 4, Appli |
| 39 | 83 | 59.7 | 466 | 16 | US-09-267-574-2 | Sequence 2, Appli |
| 40 | 83 | 59.7 | 466 | 16 | US-09-267-574-4 | Sequence 4, Appli |
| 41 | 83 | 59.7 | 466 | 18 | US-09-409-918-2 | Sequence 2, Appli |
| 42 | 83 | 59.7 | 466 | 18 | US-09-409-918-4 | Sequence 4, Appli |
| 43 | 82 | 59.0 | 116 | 24 | US-60-134-600-144 | Sequence 144, App |
| 44 | 82 | 59.0 | 116 | 24 | US-60-140-956-1924 | Sequence 1924, Ap |
| 45 | 82 | 59.0 | 154 | 24 | US-60-146-394-963 | Sequence 963, App |

ALIGNMENTS

RESULT 1

US-08-089-989-1
; Sequence 1, Application US/08089989
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.
; APPLICANT: JOHANSEN, JULIA S.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/089,989
; FILING DATE: 09-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: PD2759
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110

; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: YKL-40 N-TERMINAL SEQUENCE
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..25
; US-08-089-989-1

Query Match 98.6%; Score 137; DB 4; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.9e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKLVYITWSQYREGDGSXFPDAL 25
|||||
DB 1 YKLVYITWSQYREGDGSXFPDAL 25

RESULT 2

US-09-164-862B-1
; Sequence 1, Application US/09164862B
; GENERAL INFORMATION:
; APPLICANT: Price, Paul
; APPLICANT: Johansen, Julia
; TITLE OF INVENTION: YKL-40 AS A MARKER AND PROGNOSTIC INDICATOR FOR CANCERS
; FILE REFERENCE: 407T-8955-0005
; CURRENT APPLICATION NUMBER: US/09/164,862B
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens.
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (20)..(20)
; OTHER INFORMATION: Residue unidentified
US-09-164-862B-1

Query Match 98.6%; Score 137; DB 15; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.9e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKLVYITWSQYREGDGSXFPDAL 25
|||||
DB 1 YKLVYITWSQYREGDGSXFPDAL 25

RESULT 3

US-09-215-077A-1
; Sequence 1, Application US/09215077A
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.
; APPLICANT: JOHANSEN, JULIA S.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; FILE REFERENCE: 407T-895411US
; CURRENT APPLICATION NUMBER: US/09/215,077A
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/581,527
; PRIOR FILING DATE: 1996-04-17
; PRIOR APPLICATION NUMBER: 08/089,989
; PRIOR FILING DATE: 1993-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 1
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:N-terminal
; OTHER INFORMATION: amino acid sequence for the YKL-40 protein.
; NAME/KEY: NON_CONS
; LOCATION: (20)
; OTHER INFORMATION: X IS UNIDENTIFIED - ANY RESIDUE
US-09-215-077A-1

Query Match 98.6%; Score 137; DB 16; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.9e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 YKLVYITWSQYREGDGSXFPDAL 25
|||||
DB 1 YKLVYITWSQYREGDGSXFPDAL 25

RESULT 4

US-09-158-300-10
; Sequence 10, Application US/09158300
; GENERAL INFORMATION:
; APPLICANT: A.M.H. Boots
; APPLICANT: G.F.M. Verheijden
; APPLICANT: E.S. Bos
; TITLE OF INVENTION: Novel Peptides derived from autoantigen for use
; TITLE OF INVENTION: in Immunotherapy of Autoimmune Diseases
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Akzo Nobel Patent Department
; STREET: 1300 Piccard Drive, Suite 206
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/158,300
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/634,493
; FILING DATE: 18-APR-1996
; APPLICATION NUMBER: US 08/619,645
; FILING DATE: 25-MAR-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP95/04201
; FILING DATE: 25-OCT-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL 942031287
; FILING DATE: 27-OCT-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: NL 952008860
; FILING DATE: 07-APR-1995
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 258-5200
; TELEFAX: (301) 977-0847
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 43 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-09-158-300-10

Query Match 97.8%; Score 136; DB 15; Length 43;
Best Local Similarity 96.0%; Pred. No. 4.6e-12;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKLVYITWSQYREGDGSXFFDAL 25
|||||
Db 1 YKLVYITWSQYREGDGSXFFDAL 25

RESULT 5

US-09-171-562-1
; Sequence 1, Application US/09171562
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Glibertus F.M.
; APPLICANT: Bos, Ebo S.
; TITLE OF INVENTION: Autoantigen and Proteins Structurally Related thereto
; TITLE OF INVENTION: for Use in Immunotherapy of Autoimmune Diseases
; FILE REFERENCE: O/96192US
; CURRENT FILING DATE: 1998-10-19
; EARLIER FILING DATE: 1997-04-15
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 43
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-171-562-1

Query Match 97.8%; Score 136; DB 15; Length 43;
Best Local Similarity 96.0%; Pred. No. 4.6e-12;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKLVYITWSQYREGDGSXFFDAL 25
|||||
Db 1 YKLVYITWSQYREGDGSXFFDAL 25

RESULT 6

US-09-760-475-2626
; Sequence 2626, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT249
; CURRENT APPLICATION NUMBER: US/09760,475
; PRIOR FILING DATE: 2001-01-16
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2626
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-475-2626

Query Match 97.8%; Score 136; DB 21; Length 206;
Best Local Similarity 96.0%; Pred. No. 2.2e-11;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKLVYITWSQYREGDGSXFFDAL 25
|||||
Db 43 YKLVYITWSQYREGDGSXFFDAL 67

RESULT 7

PCT-US01-14827-14282
; Sequence 14282, Application PC/TUS0114827

; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR FILING DATE: 09/577,408
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 14282
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (1)..(336)
; OTHER INFORMATION: Glycosyl hydrolases family 18 domain identified by Pfam,
; OTHER INFORMATION: accession name Glyco_hydro_18, E-value=9.3e-184, Pfam score of
; OTHER INFORMATION: 619.2
PCT-US01-14827-14282

Query Match 97.8%; Score 136; DB 1; Length 362;
Best Local Similarity 96.0%; Pred. No. 3.9e-11;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKLVYITWSQYREGDGSXFFDAL 25
|||||
Db 1 YKLVYITWSQYREGDGSXFFDAL 25

RESULT 8

US-08-755-051-3
; Sequence 3, Application US/08755051
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Au-Young, Janice
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CHITOTRIOSIDASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,051
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0152 US
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 383 amino acids
; TYPE: amino acid
; STRANDEDNESS: single

Db 32 YQLTCYFTNWAQYRPGG

```

; LENGTH: 383
;
; TYPE: PRT
;
; ORGANISM: Sus scrofa
;

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:21:26 ; Search time 169,18 Seconds
(without alignments)
1.495 Million cell updates/sec

Title: US-09-164-862B-1

Perfect score: 139

Sequence: 1 YKLVCYITWSQYREGDSXFPDAL 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 56692 seqs, 10114640 residues

Total number of hits satisfying chosen parameters: 56692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 78 | 56.1 | 423 | 4 | US-08-850-348-2 |
| 2 | 50 | 36.0 | 437 | 5 | US-09-815-242-5664 |
| 3 | 50 | 36.0 | 520 | 5 | US-09-815-242-12205 |
| 4 | 43 | 30.9 | 570 | 5 | US-09-877-633-13 |
| 5 | 43 | 30.9 | 570 | 5 | US-09-877-633-14 |
| 6 | 42 | 30.2 | 365 | 5 | US-09-620-394B-2582 |
| 7 | 42 | 30.2 | 452 | 5 | US-09-760-446A-1771 |
| 8 | 41 | 29.5 | 795 | 1 | PCT-US01-27760-855 |
| 9 | 41 | 29.5 | 802 | 5 | US-09-978-189-169 |
| 10 | 41 | 29.5 | 802 | 5 | US-09-978-189-169 |
| 11 | 41 | 29.5 | 802 | 5 | US-09-978-189-169 |
| 12 | 41 | 29.5 | 802 | 5 | US-09-978-697-169 |
| 13 | 41 | 29.5 | 802 | 5 | US-09-978-824-169 |
| 14 | 41 | 29.5 | 802 | 5 | US-09-888-615-113 |
| 15 | 40 | 28.8 | 229 | 5 | US-09-974-449-37 |
| 16 | 40 | 28.8 | 364 | 5 | US-09-978-189-515 |
| 17 | 40 | 28.8 | 364 | 5 | US-09-978-189-515 |
| 18 | 40 | 28.8 | 364 | 5 | US-09-978-189-515 |
| 19 | 40 | 28.8 | 364 | 5 | US-09-978-824-515 |
| 20 | 39.5 | 28.4 | 348 | 5 | US-09-815-242-10718 |
| 21 | 39 | 28.1 | 162 | 5 | US-09-663-600A-126 |
| 22 | 39 | 28.1 | 162 | 5 | US-09-663-600A-220 |
| 23 | 39 | 28.1 | 216 | 5 | US-09-826-734-154 |
| 24 | 39 | 28.1 | 269 | 5 | US-09-965-212-26 |
| 25 | 39 | 28.1 | 269 | 5 | US-09-965-212-28 |
| 26 | 39 | 28.1 | 328 | 5 | US-09-888-615-91 |
| 27 | 39 | 28.1 | 413 | 5 | US-09-976-594-562 |

28 39 28.1 454 5 US-09-888-615-103 Sequence 103, App
29 39 28.1 461 5 US-09-989-723-314 Sequence 314, App
30 39 28.1 461 5 US-09-989-724-314 Sequence 314, App
31 39 28.1 461 5 US-09-989-730-314 Sequence 314, App
32 39 28.1 461 5 US-09-990-446-314 Sequence 314, App
33 39 28.1 461 5 US-09-989-721-314 Sequence 314, App
34 39 28.1 461 5 US-09-989-722-314 Sequence 314, App
35 39 28.1 461 5 US-09-989-723-314 Sequence 314, App
36 39 28.1 461 5 US-09-989-725-314 Sequence 314, App
37 39 28.1 461 5 US-09-989-726-314 Sequence 314, App
38 39 28.1 461 5 US-09-989-727-314 Sequence 314, App
39 39 28.1 461 5 US-09-989-728-314 Sequence 314, App
40 39 28.1 461 5 US-09-989-731-314 Sequence 314, App
41 39 28.1 461 5 US-09-989-734-314 Sequence 314, App
42 39 28.1 461 5 US-09-989-734-314 Sequence 314, App
43 39 28.1 461 5 US-09-990-437-314 Sequence 314, App
44 39 28.1 461 5 US-09-990-438-314 Sequence 314, App
45 39 28.1 461 5 US-09-990-440-314 Sequence 314, App

ALIGNMENTS

RESULT 1
US-08-850-348-2
; Sequence 2, Application US/08850348
; GENERAL INFORMATION:
; APPLICANT: Kirkpatrick, Robert
; Rosenberg, Martin
; TITLE OF INVENTION: Human Cartilage Glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/850,348
; FILING DATE: 02-May-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,532
; FILING DATE: 03-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 5219
; REFERENCE/DOCKET NUMBER: P50390
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-850-348-2

Query Match 56.1%; Score 78; DB 4; Length 423;
Best Local Similarity 56.5%; Pred. No. 0.00015;
Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

```
QY 1 YKLVYVTSWSQYREGDGSXFPD 23
      |||||:|||||:| |
Db 60 YKLVYFTNKSQDRQPGKFTPE 82

RESULT 2
US-09-815-242-5664
; Sequence 5664, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5664
; LENGTH: 437
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5664

Query Match 36.0%; Score 50; DB 5; Length 437;
Best Local Similarity 40.0%; Pred. No. 2;
Matches 10; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

QY 1 YKLVYVTSWSQYREGDGSXFPD 25
      |||||:|||||:| |
Db 179 YVMVCDTAWSYMGQDNDFKDQL 203

RESULT 3
US-09-815-242-12205
; Sequence 12205, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21

QY 1 YKLVYVTSWSQYREGDGSXFPD 25
      |||||:|||||:| |
Db 179 YVMVCDTAWSYMGQDNDFKDQL 203

RESULT 4
US-09-877-633-13
; Sequence 13, Application US/09877633
; GENERAL INFORMATION:
; APPLICANT: Jenniffer Hillman
; TITLE OF INVENTION: DIAGNOSTIC MARKER FOR CANCERS
; FILE REFERENCE: PC-0040 CIP
; CURRENT APPLICATION NUMBER: US/09/877,633
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; OTHER INFORMATION: Incyte ID No: g6449083
US-09-877-633-13

Query Match 30.9%; Score 43; DB 5; Length 570;
Best Local Similarity 40.0%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 YVTSWSQYREGDGSXFPD 25
      |||||:|||||:| |
Db 85 YLTSYGQLSNGEPHFLPDAM 104

RESULT 5
US-09-877-633-14
; Sequence 14, Application US/09877633
; GENERAL INFORMATION:
; APPLICANT: Jenniffer Hillman
; TITLE OF INVENTION: DIAGNOSTIC MARKER FOR CANCERS
; FILE REFERENCE: PC-0040 CIP
; CURRENT APPLICATION NUMBER: US/09/877,633
; CURRENT FILING DATE: 2001-06-08
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PERL Program
US-09-877-633-13
```

```
; SEQ ID NO 14
; LENGTH: 570
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: g5360085
US-09-877-633-14

Query Match          30.9%; Score 43; DB 5; Length 570;
Best Local Similarity 40.0%; Pred. No. 27;
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 6 YVTSWYREGDGSXFPDAL 25
   |||:| | | | | | |
Db 85 YLTSYQLSNGEPHFLPDAM 104

RESULT 6
US-09-620-394B-2582
; Sequence 2582, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1087P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 2582
; LENGTH: 365
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..365
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..365
; OTHER INFORMATION: Ceres Seq. ID 1382327
US-09-620-394B-2582

Query Match          30.2%; Score 42; DB 5; Length 365;
Best Local Similarity 42.9%; Pred. No. 25;
Matches 9; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 2 KLVCIYTSWYREGDGSXFP 22
   :|| :|| :|| :|| :||
Db 45 KLVVRSSGSDYQNGDVNGFP 65

RESULT 7
US-09-760-446A-1771
; Sequence 1771, Application US/09760446A
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT268
; CURRENT APPLICATION NUMBER: US/09/760,446A
; CURRENT FILING DATE: 2000-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/225,757
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/226,868
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/216,647
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,267
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/216,880
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: 60/225,270
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/251,869
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/235,834
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/234,274
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/234,223
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: 60/228,924
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/224,518
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,369
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/224,519
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,964
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/241,809
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/249,299
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/236,327
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/241,785
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/244,617
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/225,268
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/236,368
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/251,856
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/251,868
; PRIOR FILING DATE: 2000-12-08
; PRIOR APPLICATION NUMBER: 60/229,344
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/234,997
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: 60/229,343
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,345
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,287
; PRIOR FILING DATE: 2000-09-01
; PRIOR APPLICATION NUMBER: 60/229,513
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/231,413
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/229,509
; PRIOR FILING DATE: 2000-09-05
; PRIOR APPLICATION NUMBER: 60/236,367
```

;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/237,039
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,038
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/236,370
;; PRIOR FILING DATE: 2000-09-29
;; PRIOR APPLICATION NUMBER: 60/236,802
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,037
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/237,040
;; PRIOR FILING DATE: 2000-10-02
;; PRIOR APPLICATION NUMBER: 60/240,960
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/239,935
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/239,937
;; PRIOR FILING DATE: 2000-10-13
;; PRIOR APPLICATION NUMBER: 60/241,787
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,474
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/246,532
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/249,216
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,210
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/226,681
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,759
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/225,213
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/227,182
;; PRIOR FILING DATE: 2000-08-22
;; PRIOR APPLICATION NUMBER: 60/225,214
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/235,836
;; PRIOR FILING DATE: 2000-09-27
;; PRIOR APPLICATION NUMBER: 60/230,438
;; PRIOR FILING DATE: 2000-09-06
;; PRIOR APPLICATION NUMBER: 60/215,135
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: 60/225,266
;; PRIOR FILING DATE: 2000-08-14
;; PRIOR APPLICATION NUMBER: 60/249,218
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,208
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,213
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,212
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,207
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,245
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,244
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,217
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,211
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,215
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,264
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,214
;; PRIOR FILING DATE: 2000-11-17
;; PRIOR APPLICATION NUMBER: 60/249,297
;; PRIOR FILING DATE: 2000-11-17

;; PRIOR APPLICATION NUMBER: 60/232,400
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/231,242
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,081
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/232,080
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,414
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/231,244
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,064
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/233,063
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,397
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,399
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,401
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/241,808
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,826
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,786
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/241,221
;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/246,475
;; PRIOR FILING DATE: 2000-11-08
;; PRIOR APPLICATION NUMBER: 60/231,243
;; PRIOR FILING DATE: 2000-09-08
;; PRIOR APPLICATION NUMBER: 60/233,065
;; PRIOR FILING DATE: 2000-09-14
;; PRIOR APPLICATION NUMBER: 60/232,398

Query Match 30.2%; Score 42; DB 5; Length 452;
Best Local Similarity 58.3%; Pred. No. 30;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 13 YREGDGSXFPDA 24
| | | : | | | |
Db 160 YNTGDNTPPDA 171

RESULT 8

US-09-611-451-76
; Sequence 76, Application US/09611451
; GENERAL INFORMATION:
; APPLICANT: Barbas, Carlos F.
; Burton, Dennis R.
; Lerner, Righard A.
; TITLE OF INVENTION: SYNTHETIC HUMAN NEUTRALIZING MONOCLONAL
; ANTIBODIES TO HUMAN IMMUNODEFICIENCY VIRUS
; NUMBER OF SEQUENCES: 92
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: The Scripps Research Institute, Office of
; Patent Counsel
; STREET: 10550 North Torrey Pines Road, TPC 8
; CITY: La Jolla
; STATE: CA
; COUNTRY: USA
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0
; CURRENT APPLICATION DATA: US/09/611,451
; APPLICATION NUMBER: 06-Jul-2000
; FILING DATE: 06-Jul-2000

PRIOR APPLICATION DATA:
APPLICANT: Ferrara, Napoleon
FILING DATE: 2001-10-29
APPLICATION NUMBER: 08/591,632
FILING DATE: 2001-10-29
APPLICATION NUMBER: US 08/308,841
FILING DATE: 19-SEP-1994
APPLICATION NUMBER: US 08/233,619
FILING DATE: 26-APR-1994
APPLICATION NUMBER: US 08/139,409
FILING DATE: 19-OCT-1993
ATTORNEY/AGENT INFORMATION:
NAME: Fitting, Thomas
REGISTRATION NUMBER: 34,163
REFERENCE/DOCKET NUMBER: TSRI 332.3
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 784-2937
TELEFAX: (619) 784-9399
INFORMATION FOR SEQ ID NO: 76:
SEQUENCE CHARACTERISTICS:
LENGTH: 108 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: Protein
SEQUENCE DESCRIPTION: SEQ ID NO: 76:
US-09-611-451-76

Query Match 29.5%; Score 41; DB 5; Length 108;
Best Local Similarity 46.7%; Pred. No. 11;
Matches 7; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 5 CYYTSMQYREGDVS 19
| | | | | | | | | |
Db 87 CQYGVHSQYTFGGGT 101

RESULT 9
PCT-US01-27760-855
Sequence 855, Application PC/TUS0127760
GENERAL INFORMATION:
APPLICANT: Hyseq, Inc
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 21272-115/ 785
CURRENT APPLICATION NUMBER: PCT/US01/27760
CURRENT FILING DATE: 2001-10-11
PRIOR APPLICATION NUMBER: 09/687,527
PRIOR FILING DATE: 2000-10-12
NUMBER OF SEQ ID NOS: 892
SOFTWARE: Custom
SEQ ID NO 855
LENGTH: 795
TYPE: PRT
ORGANISM: Homo sapiens
PCT-US01-27760-855

Query Match 29.5%; Score 41; DB 1; Length 795;
Best Local Similarity 50.0%; Pred. No. 72;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 LVCYYTSMQYREG 16
| | | | | | | | | |
Db 684 LHCWITGNGALREG 697

RESULT 10
US-09-978-189-169
Sequence 169, Application US/09978189
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan

APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630P1C7
CURRENT APPLICATION NUMBER: US/09/978,189
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30

;; PRIOR APPLICATION NUMBER: 60/079923
;; PRIOR FILING DATE: 1998-03-30
;; PRIOR APPLICATION NUMBER: 60/080105
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080107
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080165
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080194
;; PRIOR FILING DATE: 1998-03-31
;; PRIOR APPLICATION NUMBER: 60/080327
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080328
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080333
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/080334
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/081070
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081049
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081071
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081195
;; PRIOR FILING DATE: 1998-04-08
;; PRIOR APPLICATION NUMBER: 60/081203
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081229
;; PRIOR FILING DATE: 1998-04-09
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081817
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081952
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081838
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082568
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082569
;; PRIOR FILING DATE: 1998-04-21
;; PRIOR APPLICATION NUMBER: 60/082704
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082700
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082797
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/082796
;; PRIOR FILING DATE: 1998-04-23
;; PRIOR APPLICATION NUMBER: 60/083336
;; PRIOR FILING DATE: 1998-04-27
;; PRIOR APPLICATION NUMBER: 60/083322
;; PRIOR FILING DATE: 1998-04-28
;; PRIOR APPLICATION NUMBER: 60/083392
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083495
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083496
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083499
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083545
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083554
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083558
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083559

;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083500
;; PRIOR FILING DATE: 1998-04-29
;; PRIOR APPLICATION NUMBER: 60/083742
;; PRIOR FILING DATE: 1998-04-30
;; PRIOR APPLICATION NUMBER: 60/084366
;; PRIOR FILING DATE: 1998-05-05
;; PRIOR APPLICATION NUMBER: 60/084414
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/084637
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084639
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084640
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084598
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084600
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084627
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/084643
;; PRIOR FILING DATE: 1998-05-07
;; PRIOR APPLICATION NUMBER: 60/085339
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085338
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085582
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085700
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085689
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085580
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085573
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085704
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/085697
;; PRIOR FILING DATE: 1998-05-15

Query Match 29.5%; Score 41; DB 5; Length 802;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 LVCYVTSWSQYREG 16

Db 691 LHCWITGWLREG 704

RESULT 11

US-09-978-192-169
; Sequence 169, Application US/09978192
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey

APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kijavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.
APPLICANT: Pan, James.
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: P2630PIC9
CURRENT APPLICATION NUMBER: US/09/978,192
CURRENT FILING DATE: 2001-10-15
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05

APPLICANT: Pan, James;
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Shelton, David L.
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC27
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079786
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01

Query Match 29.5%; Score 41; DB 5; Length 802;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 LVCVYTSWSQVREG 16
| | | | |
Db 691 LHCWITGNGALREG 704

RESULT 12
US-09-978-697-169
Sequence 169, Application US/09978697
GENERAL INFORMATION:
APPLICANT: Ashkenazi, Avi
APPLICANT: Baker Kevin P.
APPLICANT: Botstein, David
APPLICANT: Desnoyers, Luc
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleon
APPLICANT: Filvaroff, Ellen
APPLICANT: Fong, Sherman
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerber, Hanspeter
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Hillan, Kenneth J.
APPLICANT: Kljavin, Ivar J.
APPLICANT: Kuo, Sophia S.
APPLICANT: Napier, Mary A.

;
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639

;
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15

Query Match 29.5%; Score 41; DB 5; Length 802;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 LVCYVTSWSQYREG 16

Db 691 LHCWITGWGALREG 704

RESULT 13

US-09-978-824-169
; Sequence 169, Application US/09978824

; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J.
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey

APPLICANT: Wood, William I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
FILE REFERENCE: P2630PIC14
CURRENT APPLICATION NUMBER: US/09/978,824
PRIOR FILING DATE: 2001-10-17
PRIOR APPLICATION NUMBER: 09/918585
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064249
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/065311
PRIOR FILING DATE: 1997-11-13
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/077450
PRIOR FILING DATE: 1998-03-10
PRIOR APPLICATION NUMBER: 60/077632
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077641
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077649
PRIOR FILING DATE: 1998-03-11
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078004
PRIOR FILING DATE: 1998-03-13
PRIOR APPLICATION NUMBER: 60/078886
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078936
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/078939
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079664
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079689
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/079920
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/079923
PRIOR FILING DATE: 1998-03-30
PRIOR APPLICATION NUMBER: 60/080105
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080107
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080194
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/080327
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080328
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080333
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/080334
PRIOR FILING DATE: 1998-04-01
PRIOR APPLICATION NUMBER: 60/081070
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081049
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081071
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081195
PRIOR FILING DATE: 1998-04-08
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081952
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081838
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082568
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082569
PRIOR FILING DATE: 1998-04-21
PRIOR APPLICATION NUMBER: 60/082704
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082700
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082797
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/082796
PRIOR FILING DATE: 1998-04-23
PRIOR APPLICATION NUMBER: 60/083336
PRIOR FILING DATE: 1998-04-27
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083392
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083495
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083496
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083499
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083554
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083558
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083559
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083500
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/083742
PRIOR FILING DATE: 1998-04-30
PRIOR APPLICATION NUMBER: 60/084366
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/084414
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084639
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084640
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-5-07

; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697
; PRIOR FILING DATE: 1998-05-15

Query Match 29.5%; Score 41; DB 5; Length 802;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 LVCYVTSWSQYREG 16
| | | | |
Db 691 LHCWITGWAALREG 704

RESULT 14
US-09-888-615-113
; Sequence 113, Application US/09888615
; GENERAL INFORMATION:
; APPLICANT: PLOWMAN, GREGORY
; APPLICANT: WHYTE, DAVID
; APPLICANT: CAENEPEEL, SEAN
; APPLICANT: CHARYDCZAK, GLEN
; APPLICANT: MANNING, GERARD
; APPLICANT: SUDARSANAM, SUCHA
; TITLE OF INVENTION: NOVEL PROTEASES
; FILE REFERENCE: 038602/1214
; CURRENT APPLICATION NUMBER: US/09/888,615
; CURRENT FILING DATE: 2001-06-26
; PRIOR APPLICATION NUMBER: 60/214,047
; PRIOR FILING DATE: 2000-06-26
; NUMBER OF SEQ ID NOS: 150
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 113
; LENGTH: 802
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-888-615-113

Query Match 29.5%; Score 41; DB 5; Length 802;
Best Local Similarity 50.0%; Pred. No. 73;
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 3 LVCYVTSWSQYREG 16
| | | | |
Db 691 LHCWITGWAALREG 704

RESULT 15
US-09-974-449-37
; Sequence 37, Application US/09974449

; GENERAL INFORMATION:
; APPLICANT: Kricek, Franz
; APPLICANT: Stadler, Beda
; APPLICANT: Vogel, Monique
; TITLE OF INVENTION: ANTI-IDIOTYPIC ANTIBODIES AGAINST
; TITLE OF INVENTION: ANTIBODIES WHICH INHIBIT THE BINDING OF IMMUNOGLOBULIN E TO
; TITLE OF INVENTION: ITS HIGH AFFINITY RECEPTOR
; FILE REFERENCE: 4-30888A
; CURRENT APPLICATION NUMBER: US/09/974,449
; CURRENT FILING DATE: 2001-10-10
; PRIOR APPLICATION NUMBER: PCT/EP00/03288
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 62
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 37
; LENGTH: 229
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-974-449-37

Query Match 28.8%; Score 40; DB 5; Length 229;
Best Local Similarity 53.8%; Pred. No. 32;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 6 YVTSWSQYREGDG 18
| | | | |
Db 34 YVTSWSQYREGDG 46

Search completed: December 7, 2001, 00:35:57
Job time: 871 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 6, 2001, 23:38:40 ; Search time 71.99 Seconds
(without alignments)
26.453 Million cell updates/sec

Title: US-09-164-862b-1

Perfect score: 139

Sequence: 1 YKLVCYTSMQYREGDGSXFPDAL 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------------|
| 1 | 136 | 97.8 | 383 | 2 | A49562 cartilage glycopro |
| 2 | 134 | 96.4 | 383 | 2 | S51327 heparin-binding g1 |
| 3 | 123 | 88.5 | 36 | 2 | A27682 39k whey protein - |
| 4 | 116 | 83.5 | 405 | 2 | S61551 breast-regressing |
| 5 | 90 | 64.7 | 1635 | 2 | T14075 chitinase (EC 3.2. |
| 6 | 80 | 57.6 | 504 | 2 | A38221 chitinase (EC 3.2. |
| 7 | 70 | 50.4 | 399 | 2 | S27879 secretory glycopro |
| 8 | 70 | 50.4 | 654 | 2 | I38605 oviductal glycopro |
| 9 | 68 | 48.9 | 483 | 2 | A53918 chitinase (EC 3.2. |
| 10 | 67 | 48.2 | 539 | 2 | I46470 estrogen dependent |
| 11 | 67 | 48.2 | 617 | 2 | T15408 chitinase (EC 3.2. |
| 12 | 66 | 47.5 | 537 | 2 | S57197 oviduct-specific g |
| 13 | 59 | 42.4 | 554 | 2 | A56596 chitinase (EC 3.2. |
| 14 | 58 | 41.7 | 525 | 2 | T44445 chitinase (EC 3.2. |
| 15 | 55.5 | 39.9 | 1336 | 2 | T17479 hypothetical prote |
| 16 | 55 | 39.6 | 29 | 2 | S57204 oviduct-specific s |
| 17 | 53 | 38.1 | 315 | 2 | I38668 melanoma antigen M |
| 18 | 52 | 37.4 | 599 | 2 | D83764 chitinase BH0316 |
| 19 | 50 | 36.0 | 452 | 2 | JC4038 47k glycoprotein p |
| 20 | 50 | 36.0 | 699 | 2 | A38368 fibronogen gamma c |
| 21 | 49 | 35.3 | 447 | 2 | T16527 chitinase (EC 3.2. |
| 22 | 49 | 35.3 | 18 | 2 | S39009 hypothetical prote |
| 23 | 48 | 34.5 | 277 | 2 | S57381 oviductin - golde |
| 24 | 48 | 34.5 | 277 | 2 | S57381 protein disulfide |
| 25 | 48 | 34.5 | 1215 | 2 | T43916 chitinase A (limpor |
| 26 | 47 | 33.8 | 82 | 2 | I36924 gene MHC DQ-beta 1 |
| 27 | 47 | 33.8 | 82 | 2 | I61815 gene MHC DQ-beta 1 |
| 28 | 47 | 33.8 | 82 | 2 | I61818 gene MHC DQ-beta 1 |
| 29 | 47 | 33.8 | 82 | 2 | I61834 gene MHC DQ-beta 1 |

| | | | | | |
|----|------|------|------|---|---------------------------|
| 30 | 47 | 33.8 | 536 | 2 | G64752 xylan 1,4-beta-xy1 |
| 31 | 47 | 33.8 | 644 | 2 | T33132 hypothetical prote |
| 32 | 46.5 | 33.5 | 314 | 2 | I54519 melanoma antigen M |
| 33 | 46.5 | 33.5 | 1070 | 2 | S75712 cellulase (EC 3.2. |
| 34 | 46 | 33.1 | 204 | 2 | S03246 nef protein (clone |
| 35 | 46 | 33.1 | 209 | 2 | T01673 nef protein - huma |
| 36 | 46 | 33.1 | 531 | 2 | G83268 conserved hypotet |
| 37 | 46 | 33.1 | 783 | 2 | E65096 hypothetical 88.3k |
| 38 | 46 | 33.1 | 783 | 2 | A85869 probable isomerase |
| 39 | 46 | 33.1 | 975 | 2 | T30816 macrophage colony- |
| 40 | 45.5 | 32.7 | 884 | 2 | T18649 hypothetical prote |
| 41 | 45 | 32.4 | 108 | 2 | D72617 hypothetical prote |
| 42 | 45 | 32.4 | 142 | 2 | S72302 ribosomal protein |
| 43 | 45 | 32.4 | 205 | 1 | ASLJIK nef protein - siml |
| 44 | 45 | 32.4 | 297 | 2 | T06500 alpha/beta-gliadin |
| 45 | 45 | 32.4 | 326 | 2 | D22364 alpha/beta-gliadin |

ALIGNMENTS

```

RESULT 1
A49562
cartilage glycoprotein gp39 precursor - human
N:Alternate names: 39k synovial protein
C:Species: Homo sapiens (man)
C>Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999
C:Accession: A49562; S10677; A33162
R:Hakala, B.E.; White, C.; Recklies, A.D.
J. Biol. Chem. 268, 25803-25810, 1993
A>Title: Human cartilage gp-39, a major secretory product of articular chondrocytes a
A:Reference number: A49562; M01D:94064658
A:Accession: A49562
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-383 <HAK>
A:Cross-references: GB:M80927; NID:g348911; PIDN:AA16074.1; PID:g348912
R:Mytkos, P.; Golds, F.E.
Biochem. J. 269, 265-268, 1990
A>Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary pr
A:Reference number: S10677; M01D:90328983
A:Accession: S10677
A:Molecule type: protein
A:Residues: 22-40,'X',42-45 <NY2>
C:Superfamily: Streptomyces chitinase chl40
C:Keywords: cartilage; extracellular protein; glycoprotein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 97.8%; Score 136; DB 2; Length 383;
Best Local Similarity 96.0%; Pred. No. 1,1e-11;
Matches 24; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKLVCYTSMQYREGDGSXFPDAL 25
DB 22 YKLVCYTSMQYREGDGSXFPDAL 46

RESULT 2
S51327
heparin-binding glycoprotein 38k - pig
C:Species: Sus scrofa domestica (domestic pig)
C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999
R:Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.
submitted to the EMBL Data Library, January 1995
A:Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in diff
A:Reference number: S51327
A:Accession: S51327
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-383 <SHA>

```

A:Cross-references: EMBL:Z47803; NID:g634097; PIDN:CAA87764.1; PID:g634098
C:Superfamily: Streptomyces chitinase chi40

Query Match 96.4%; Score 134; DB 2; Length 383;
Best Local Similarity 92.0%; Pred. No. 2e-11;
Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 YKLVGYTSMQYREGDSXFPDAI 25
|||||:|||||:|||||:|||||:
Db 22 YKLVGYTSMQYREGDSXFPDAI 46

RESULT 3
A27682
39k whey protein - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 21-Aug-1998
C:Accession: A27682
R:Rejman, J.J.; Hurley, W.L.
Biochem. Biophys. Res. Commun. 150, 329-334, 1988
A:Title: Isolation and characterization of a novel 39 kilodalton whey protein from bovin

A:Reference number: A27682; MUID:8810603
A:Accession: A27682
A:Molecule type: protein
A:Residues: 1-36 <REU>
C:Superfamily: Streptomyces chitinase chi40

Query Match 88.5%; Score 123; DB 2; Length 36;
Best Local Similarity 88.0%; Pred. No. 8.5e-11;
Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 YKLVGYTSMQYREGDSXFPDAI 25
|||:|||||:|||||:|||||:
Db 1 YKLVGYTSMQYREGDSXFPDAI 25

RESULT 4
S61551
breast-regressing protein brip3 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S61551; S61550; I48271
R:Morrison, B.W.; Leder, P.
Oncogene 9, 3417-3426, 1994
A:Title: new and ras initiate murine mammary tumors that share genetic markers generally
A:Reference number: I48271; MUID:95060797
A:Accession: S61551
A>Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-405 <MOR1>
A:Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
R:Morrison, B.W.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61550

A:Accession: S61550
A:Molecule type: mRNA
A:Residues: 1-245, 'I', 247-330, 'H', 332-350, 'MYMALDLDFQGTGCPKKEFFLTNAIKDALA' <MOR2>
A:Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066
A>Note: the differences at the carboxyl end are due to a frameshift error
C:Genetics:
A:Gene: brip3
C:Superfamily: Streptomyces chitinase chi40
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-405/Product: breast-regressing protein brip3 #status predicted <MAT>

Query Match 83.5%; Score 116; DB 2; Length 405;
Best Local Similarity 80.0%; Pred. No. 7.3e-09;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YKLVGYTSMQYREGDSXFPDAI 25

Db 22 YKLVGYTSMQYREGDSXFPDAI 46
|||||:|||||:|||||:|||||:

RESULT 5
T14075
chitinase (EC 3.2.1.14) - yellow fever mosquito
C:Species: Aedes aegypti (yellow fever mosquito)
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
C:Accession: T14075
R:de la Vega, H.; Specht, C.A.; Liu, Y.; Robbins, P.W.
Insect Mol. Biol. 7, 233-239, 1997
A:Title: Chitinases are a multi-gene family in Aedes, Anopheles, and Drosophila.

A:Accession: T14075
A:Reference number: Z17872
A:Molecule type: DNA
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Residues: 1-1635
A:Cross-references: EMBL:AF026492; NID:g2564720; PID:g2564721; PIDN:AA81850.1
C:Genetics:
A:Gene: CHP2
A:Introns: 462/3; 524/3; 618/1; 951/3; 1151/2
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 64.7%; Score 90; DB 2; Length 1635;
Best Local Similarity 56.5%; Pred. No. 0.00012;
Matches 13; Conservative 6; Mismatches 4; Indels 0; Gaps 0;

OY 1 YKLVGYTSMQYREGDSXFPD 23
|||:|||||:|||||:|||||:
Db 1257 FKVVGYFTNMAWYRGDSKYPD 1279

RESULT 6
A38221
chitinase (EC 3.2.1.14) MFL - nematode (Brugia malayi)
C:Species: Brugia malayi
C>Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000
C:Accession: A38221
R:Fuhrman, J.A.; Lane, W.S.; Smith, R.F.; Plessens, W.F.; Perler, F.B.
Proc. Natl. Acad. Sci. U.S.A. 89, 1548-1552, 1992
A:Title: Transmision-blocking antibodies recognize microfilarial chitinase in Brugia
A:Reference number: A38221; MUID:92179220
A:Accession: A38221
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid; protein
A:Residues: 1-504 <PUH>
A:Cross-references: GB:M73689; NID:g156063; PIDN:AA27854.1; PID:g156064
A>Note: sequence extracted from NCBI backbone (NCBI:05345)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 57.6%; Score 80; DB 2; Length 504;
Best Local Similarity 61.1%; Pred. No. 0.001;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

OY 5 CYTWSQYREGDSXFPD 22
|||||:|||||:|||||:|||||:
Db 27 CYTNMAQYRGDSKFLP 44

RESULT 7
S27879
secretory protein YM-1 precursor - mouse
C:Species: Mus musculus (house mouse)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 22-Jun-1999
C:Accession: S27879
R:Chang, N.C.A.; Liu, C.H.; Chang, A.C.
submitted to the EMBL Data Library, June 1992
A:Description: Molecular characterization of a secretory protein (YM-1) transiently e
A:Reference number: S27879
A:Accession: S27879

A:Molecule type: mRNA
A:Residues: 1-399 <CH>
A:Cross-references: EMBL:M94584; NID:g202441; PIDN:AAB62394.1; PID:g202442
C:Superfamily: Streptomyces chitinase chl40
F:1-21/Domain: signal sequence status predicted <SIG>
F:22-399/Product: secretory protein YM-1 #status predicted <MAN>

Query Match 50.4%; Score 70; DB 2; Length 399;
Best Local Similarity 54.5%; Pred. No. 0.022;
Matches 12; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 YKLVCTYSMSQYREGDSXFP 22
:|||||:|:|:|
Db 22 YOLMCYTTSMADRPLEGSEFKP 43

RESULT 8
138605

oviductal glycoprotein - human
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: 138605
R:Atlas, E.B.; Verhage, H.G.; Jaffe, R.C.
Biol. Reprod. 51, 685-694, 1994
A>Title: Complementary deoxyribonucleic acid cloning and molecular characterization of a
A:Reference number: 138605; MUID:95119256
A:Accession: 138605
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-654 <RES>
A:Cross-references: EMBL:U09550; NID:g529147; PID:g529148

Query Match 50.4%; Score 70; DB 2; Length 654;
Best Local Similarity 48.0%; Pred. No. 0.034;
Matches 12; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 YKLVCTYSMSQYREGDSXFPAL 25
:|||||:|:|:|
Db 22 HRLVCYFTWMAHSRPGPASILPHDL 46

RESULT 9
A53918

chitinase (EC 3.2.1.14) precursor - braconid wasp (Chelonus sp.)
C:Species: Chelonus sp.
C>Date: 28-Jul-1995 #sequence_revision 28-Jul-1995 #text_change 21-Jul-2000
C:Accession: A53918
R:Krishnan, A.; Nair, P.N.; Jones, D.
J. Biol. Chem. 269, 20971-20976, 1994
A>Title: Isolation, cloning, and characterization of new chitinase stored in active form
A:Reference number: A53918; MUID:94342256
A:Accession: A53918
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-483 <KR>
A:Cross-references: GB:U10422; NID:g53504; PIDN:AAA61639.1; PID:g53505
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 48.9%; Score 68; DB 2; Length 483;
Best Local Similarity 58.8%; Pred. No. 0.045;
Matches 10; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

OY 2 KLVCTYSMSQYREGDG 18
:||||:|:|:|:|
Db 23 KVCYCGAMSVYRGNG 39

RESULT 10
16470

estrogen dependent oviduct protein precursor - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 03-Nov-1999
C:Accession: I46470
R:Desouza, M.M.; Murray, M.K.
Endocrinology 136, 2485-2496, 1995
A>Title: An estrogen-dependent secretory protein, which shares identity with chitinas
and embryo development

A:Reference number: I46470; MUID:95269691
A:Accession: I46470
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-539 <DES>
A:Cross-references: EMBL:U16719; NID:g885600; PIDN:AAC48471.1; PID:g885601

Query Match 48.2%; Score 67; DB 2; Length 539;
Best Local Similarity 50.0%; Pred. No. 0.075;
Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

OY 1 YKLVCTYSMSQYREGDSXFP 22
:|||||:|:|:|
Db 22 HRLVCYFTWMAHSRPGSASTLP 43

RESULT 11
T15408

hypothetical protein C04F6.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 20-Sep-1999
R:Nhan, M.
submitted to the EMBL Data Library, December 1995
A:Description: The sequence of C. elegans cosmid C04F6.
A:Reference number: T15408
A:Accession: T15408
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-617 <NHA>
A:Cross-references: EMBL:U42835; NID:g1125760; PID:g1125762; PIDN:AAA83586.1; CESP:C0
A:Genetics:
A:Gene: CESP:C04F6.3
A:Introns: 28/1; 66/2; 504/1

Query Match 48.2%; Score 67; DB 2; Length 617;
Best Local Similarity 47.4%; Pred. No. 0.085;
Matches 9; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 5 CYTSMQYREGDSXFPD 23
||:|:|:|:|:|
Db 57 CYPTNMAQYRQGRKRVPE 75

RESULT 12
S57197

oviduct-specific glycoprotein 95K precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C>Date: 28-Oct-1995 #sequence_revision 03-Nov-1995 #text_change 05-Nov-1999
C:Accession: S57197
R:Sendai, Y.; Abe, H.; Kikuchi, M.; Satoh, T.; Hoshi, H.
Biol. Reprod. 50, 927-934, 1994
A>Title: Purification and molecular cloning of bovine oviduct-specific glycoprotein.
A:Reference number: S57197; MUID:94257768
A:Accession: S57197
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-537 <SEN>
A:Cross-references: EMBL:D16639; NID:g391621; PIDN:BAA04065.1; PID:d1004583; PID:g391
C:Keywords: glycoprotein

Query Match 47.5%; Score 66; DB 2; Length 537;
Best Local Similarity 50.0%; Pred. No. 0.1;

Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
QY 1 YKLVCTYSMSQYREGDSXFP 22
:|||||:|:| | | |
Db 19 HKLVCTFTNMAFSRPGASITLP 40

RESULT 13

A56596
chitinase (EC 3.2.1.14) - tobacco hornworm
C:Species: Manduca sexta (tobacco hornworm)
C:Date: 11-Aug-1995 #sequence_revision 11-Aug-1995 #text_change 29-Jan-1999
R:Kramer, K.D.; Corpuz, L.; Choi, H.K.; Muthukrishnan, S.
Insect Biochem. Mol. Biol. 23, 691-701, 1993
A:Title: Sequence of a cDNA and expression of the gene encoding epidermal and gut chitin
A:Reference number: A56596; MID:93357793
A:Accession: A56596
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-554 <KRA>
A:Cross-references: GB:002270; GB:S64757; NID:9406048; PID:9406049
A:Experimental source: larvae
A:Note: sequence extracted from NCBI backbone (NCBIN:136417, NCBIPI:136418)
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 42.4%; Score 59; DB 2; Length 554;

Best Local Similarity 47.1%; Pred. No. 1;
Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 KLVCTYSMSQYREGDG 18
::|::|::|:| | | |
Db 24 KLVCTYSMSQYREGDG 40

RESULT 14

T44445
chitinase (EC 3.2.1.14) [imported] - African malaria mosquito
C:Species: Anopheles gambiae (African malaria mosquito)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 21-Jan-2000
C:Accession: T44445
R:Shen, Z.; Jacobs-Lorena, M.
submitted to the EMBL Data Library, June 1997
A:Reference number: 222771
A:Accession: T44445
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-525 <SHE>
A:Cross-references: EMBL:AF008575; PIDN:AAB87764.1
A:Experimental source: adult; gut
C:Genetics:
A:Gene: chl-1
C:Keywords: glycosidase; hydrolase

Query Match 41.7%; Score 58; DB 2; Length 525;

Best Local Similarity 52.9%; Pred. No. 1.4;
Matches 9; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 2 KLVCTYSMSQYREGDG 18
:|:|:|:|:|:| | | |
Db 32 KLVCTYSMSQYREGDG 48

RESULT 15

T17479
hypothetical protein PC2A361.13 - Amycolatopsis orientalis
C:Species: Amycolatopsis orientalis
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T17479
R:Van Wageningen, A.; Kirkpatrick, P.; Williams, D.; Harris, B.; Kershaw, J.; Lennard, N.
Chem.-Biol. 3, 155-162, 1998

A:Title: Sequencing and analysis of genes involved in the biosynthesis of a vancomycin
A:Reference number: Z18604
A:Accession: T17479
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1336 <VAN>
A:Cross-references: EMBL:AJ223998; NID:e1251208; PID:e1251218; PIDN:CA11770.1

Query Match 39.9%; Score 55.5; DB 2; Length 1336;

Best Local Similarity 41.7%; Pred. No. 7.1;
Matches 10; Conservative 7; Mismatches 6; Indels 1; Gaps 1;

QY 1 YKLVCTYSMSQYREGDSXFPDA 24
||:|:|:|:|:|:|:|:|
Db 614 YKLVCTYSMSQYREGDSXFPDA 636

Search completed: December 7, 2001, 00:26:03
Job time: 2843 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:24:46 ; Search time 69.03 seconds
(without alignments)
13.279 Million cell updates/sec

Title: US-09-164-862b-1

Perfect score: 139

Sequence: 1 YKLVYVTSWSYREGDSXFPDAL 25

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|---------------|---------------------|
| 1 | 136 | 97.8 | 383 | 1 C3L1_HUMAN | P36222 homo sapien |
| 2 | 123 | 88.5 | 36 | 1 C3L1_BOVIN | P30922 bos taurus |
| 3 | 116 | 83.5 | 381 | 1 C3L1_MOUSE | Q61362 mus musculus |
| 4 | 80 | 57.6 | 504 | 1 CHIT_BRUMA | P29030 brugia mala |
| 5 | 78 | 56.1 | 390 | 1 C3L2_HUMAN | Q15782 homo sapien |
| 6 | 75 | 54.0 | 721 | 1 OGP_MOUSE | Q62010 mus musculus |
| 7 | 70 | 50.4 | 623 | 1 OGP_PAPAN | P36718 papio anubi |
| 8 | 70 | 50.4 | 678 | 1 OGP_HUMAN | Q12889 homo sapien |
| 9 | 67 | 48.2 | 539 | 1 OGP_SHEEP | Q28542 ovis aries |
| 10 | 67 | 48.2 | 617 | 1 CHIT_CAEL | Q11174 caenorhabdi |
| 11 | 66 | 47.5 | 537 | 1 OGP_BOVIN | Q28042 bos taurus |
| 12 | 65 | 46.8 | 671 | 1 OGP_MESAU | Q60557 mesocricetu |
| 13 | 61 | 43.9 | 527 | 1 OGP_PIG | Q28990 sus scrofa |
| 14 | 59 | 42.4 | 554 | 1 CHIT_MANSE | P36362 manduca sex |
| 15 | 53 | 38.1 | 315 | 1 MAC9_HUMAN | P43362 homo sapien |
| 16 | 50 | 36.0 | 546 | 1 PTH2_RAT | P70555 rattus norv |
| 17 | 50 | 36.0 | 699 | 1 CHIT_BACCI | P20533 bacillus ci |
| 18 | 49 | 35.3 | 438 | 1 FIBG_XENLA | P17634 xenopus lae |
| 19 | 47 | 33.8 | 536 | 1 YAGB_ECOLI | P77713 escherichia |
| 20 | 46.5 | 33.5 | 314 | 1 MAC6_HUMAN | P43362 homo sapien |
| 21 | 46 | 33.1 | 205 | 1 NEF_HVLS3 | P19545 human immun |
| 22 | 46 | 33.1 | 208 | 1 NEF_HV1RH | P05858 human immun |
| 23 | 46 | 33.1 | 209 | 1 NEF_HV1MA | P04603 human immun |
| 24 | 46 | 33.1 | 216 | 1 NEF_HV1JR | P20867 human immun |
| 25 | 46 | 33.1 | 299 | 1 FCN3_HUMAN | O75636 homo sapien |
| 26 | 46 | 33.1 | 783 | 1 YGJK_ECOLI | P42592 escherichia |
| 27 | 45 | 32.4 | 205 | 1 NEF_SIVC2 | P17664 chimpanzee |
| 28 | 45 | 32.4 | 297 | 1 GDM4_WHEAT | P04724 triticum ae |
| 29 | 45 | 32.4 | 313 | 1 FCN2_HUMAN | Q15485 homo sapien |
| 30 | 45 | 32.4 | 326 | 1 FCN1_HUMAN | O00602 homo sapien |
| 31 | 45 | 32.4 | 550 | 1 PTHR2_HUMAN | P49190 homo sapien |
| 32 | 44.5 | 32.0 | 4543 | 1 LRPI_CHICK | P98157 gallus gall |
| 33 | 44 | 31.7 | 282 | 1 FIBA_PARPA | P19477 parastichop |

RESULT 1

| | | | | |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-----------|------|---------|
| ID | C3L1_HUMAN | STANDARD; | PRT; | 383 AA. |
| AC | P36222: P30923; | | | |
| DT | 01-JUL-1993 (Rel. 26, Created) | | | |
| DT | 01-JUN-1994 (Rel. 29, Last sequence update) | | | |
| DT | 20-AUG-2001 (Rel. 46, Last annotation update) | | | |
| DE | CHITINASE-3 LIKE PROTEIN 1 PRECURSOR (CARTILAGE GLYCOPROTEIN-39) (GP-39) (39 KDA SYNOVIAL PROTEIN) (YKL-40). | | | |
| DE | CHI3L1. | | | |
| GN | Homo sapiens (Human). | | | |
| OS | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A., AND PARTIAL SEQUENCE. | | | |
| RC | TISSUE=Cartilage; | | | |
| RX | MEDLINE=94064658; PubMed=8245017; | | | |
| RA | Hakala B.E., White C., Recklies A.D.; | | | |
| RT | "Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family."; | | | |
| RT | J. Biol. Chem. 268:25803-25810(1993). | | | |
| RN | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE=Blood; | | | |
| RX | MEDLINE=97386591; PubMed=9244440; | | | |
| RA | Rehli M., Krause S.W., Andressen R.; | | | |
| RT | "Molecular characterization of the gene for human cartilage gp-39 (CHI3L1), a member of the chitinase protein family and marker for late stages of macrophage differentiation."; | | | |
| RT | Genomics 43:221-225(1997). | | | |
| RN | [3] | | | |
| RP | SEQUENCE OF 22-45. | | | |
| RX | MEDLINE=90328983; PubMed=2375755; | | | |
| RA | Nyirkos P., Golds E.E.; | | | |
| RT | "Human synovial cells secrete a 39 kDa protein similar to a bovine mammary protein expressed during the non-lactating period."; | | | |
| RT | Biochem J 269:265-268(1990). | | | |
| CC | -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT. | | | |
| CC | -!- SUBCELLULAR LOCATION: EXTRACELLULAR. | | | |
| CC | -!- TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNOVIAL CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG, PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS. | | | |
| CC | -!- PTM: GLYCOSYLATED. | | | |
| CC | -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES. | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | |

P17801 zea mays (m
Q9v4e8 homo sapien
P25218 equine herp
P18050 equine herp
P18551 equine herp
P28922 equine herp
Q10218 schizosacch
P41614 pinus thunb
P05856 human immun
P05855 human immun
P04324 human immun
P03406 human immun

ALIGNMENTS

| | | | |
|----|---|---------------------------|----|
| QY | 1 | YKLVCYVTSWSQYREGDGSXFPDAL | 25 |
| | | : : : : | |
| Db | 1 | YKLIXYTSSWSQYREGDGSXFPDAI | 25 |

RESULT 3

| C3LL_MOUSE | STANDARD; | PRT; | 381 AA. |
|--------------------|----------------------------------------------------------------------------|------|-----------------------------------|
| ID | C61362; | | |
| AC | Q61362; | | |
| DT | 20-AUG-2001 (Rel. 40, Created) | | |
| DT | 20-AUG-2001 (Rel. 40, Last sequence update) | | |
| DT | 20-AUG-2001 (Rel. 40, Last annotation update) | | |
| DE | CHITINASE-3 LIKE PROTEIN 1 PRECURSOR (CARTILAGE GLYCOPROTEIN-39) | | |
| DE | (GP-39) (BRP39, PROTEIN). | | |
| GN | CHI3L1 OR BRP39. | | |
| OS | Mus musculus (Mouse). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | |
| OX | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | |
| NCBI_TaxID=10090; | | | |
| RN | [1] | | |
| SEQUENCE FROM N.A. | | | |
| RP | STRAIN=FVB/N; TISSUE=Breast; | | |
| RC | MEDLINE=95060797; PubMed=7970700; | | |
| RX | Morrison B.W., Leder P.; | | |
| RA | "neu and ras initiate murine mammary tumors that share genetic markers | | |
| RT | generally absent in c-myc and int-2-initiated tumors."; | | |
| RL | Oncogene 9:3417-3426(1994). | | |
| -1- | FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO | | |
| CC | RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY | | |
| CC | SIMILARITY). | | |
| CC | -1- SUBCELLULAR LOCATION: EXTRACELLULAR (BY SIMILARITY). | | |
| CC | -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES. | | |
| CC | ----- | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | |
| CC | use by non-profit institutions as long as its content is in no way | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | |
| CC | entities requires a license agreement (see http://www.isb-sib.ch/announce/ | | |
| CC | or send an email to license@isb-sib.ch). | | |
| CC | ----- | | |
| DR | EMBL; X93035; CAA63603.1; . | | |
| DR | MGD; MGI:1340899; Chi3l1. | | |
| DR | InterPro; IPR001579; Chitinase_2. | | |
| DR | InterPro; IPR001223; Glyco_hydro_18. | | |
| DR | Pfam; PF00704; Glyco_hydro_18; 1. | | |
| DR | PROSITE; PS01095; CHITINASE_18; FALSE_NEG. | | |
| KW | Glycoprotein; Signal. | | |
| FT | SIGNAL 1 21 | | POTENTIAL. |
| FT | CHAIN 22 381 | | CHITINASE-3 LIKE PROTEIN 1. |
| FT | CARBOHYD 60 60 | | N-LINKED (GLCNAC..) (POTENTIAL). |
| FT | SEQUENCE 381 AA; 43001 MW; EF6581E8184F0450 CRC64; | | |
| SEQ | | | |

Query Match 83.5%; Score 116; DB 1; Length 381;
Best Local Similarity 80.0%; Pred. No. 2.le-10;
Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0

QY 1 YKLVCYVTSWSQYREGDGSXFPDAL 25

|||||: |||||: |||||: |||||:

Db 1 YKLIXYTSSWSQYREGDGSXFPDAI 25

RESULT 4

| CHIT_BRUMA | STANDARD; | PRT; | 504 AA. |
|------------|------------------------------------------------------|------|---------|
| ID | C29030; | | |
| AC | P29030; | | |
| DT | 01-DEC-1992 (Rel. 24, Created) | | |
| DT | 01-DEC-1992 (Rel. 24, Last sequence update) | | |
| DT | 01-FEB-1995 (Rel. 31, Last annotation update) | | |
| DE | ENDOCHITINASE PRECURSOR (EC 3.2.1.14) (MFI ANTIGEN). | | |
| DE | Brugia malayi. | | |


```

OC Onchocercidae; Brugia.
OX NCBI_TaxID=6279;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92179220; PubMed=1542646;
RA Fuhrman J.A., Lane W.S., Smith R.F., Plessens W.F., Perler F.B.;
RT "Transmission-blocking antibodies recognize microfilarial chitinase
in brugian lymphatic filariasis.";
RL Proc. Natl. Acad. Sci. U.S.A. 89:1548-1552(1992).
CC -1- FUNCTION: THE MFI ANTIGEN IS A MICROFILARIAL CHITINASE, WHICH MAY
CC FUNCTION TO DEGRADE CHITIN-CONTAINING STRUCTURES IN THE MICRO-
CC FILARIA OR IN ITS MOSQUITO VECTOR DURING PARASITE DEVELOPMENT AND
CC TRANSMISSION.
CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF
CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -1- DEVELOPMENTAL STAGE: THE APPEARANCE OF THE MFI ANTIGEN CORRESPOND
CC WITH THE ONSET OF THE PARASITE'S ABILITY TO INFECT THE MOSQUITO.
CC -1- PTM: O-GLYCOSYLATED.
CC -1- MISCELLANEOUS: KNOWN TO BIND CALCIUM.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M73689; AAA27854.1; -
DR PIR; A38221; A38221.
DR InterPro; IPR002557; Chitin_binding.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; Chitin_bind_2; 1.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtB2; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR Hydrolase; Glycosidase; Chitin degradation; Signal; Antigen;
KW Repeat; Glycoprotein; Calcium-binding.
FT SIGNAL 1 22
FT CHAIN 23 504 ENDOCHITINASE.
FT DOMAIN 23 400 CATALYTIC.
FT DOMAIN 401 450 SER/THR-RICH (LINKER).
FT DOMAIN 407 448 3 X 14 AA APPROXIMATE TANDEM REPEATS.
FT ACT_SITE 148 148 PROTON DONOR (BY SIMILARITY).
SQ SEQUENCE 504 AA; 55971 MW; A78BE7BF8E3709B CRC64;

Query Match 57.6%; Score 80; DB 1; Length 504;
Best Local Similarity 61.1%; Pred. No. 9e-05;
Matches 11; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 5 CYYTWSQVREGDGSXFP 22
Db 27 CYYTNNQYRDGEGKFLP 44
||||:||||:|
|

RESULT 5
C3L2_HUMAN STANDARD; PRT; 390 AA.
AC Q15782; Q15783; Q15749;
DT 20-AUG-2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CHITINASE 3-LIKE PROTEIN 2 PRECURSOR (YKL-39) (CHONDROCYTE PROTEIN
DE 39).
GN CH3L2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A. (SHORT AND LONG ISOFORMS).
RA Grossman A., Matsuyama T., Baker E., Waterhouse P., Sutherland G.R.,
RA Mak T.W.;
RT "Cloning of a novel lymphoid restricted human chitinase and
RT localization to lp13.3.";
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A. (SHORT ISOFORM).
RC TISSUE=Articular cartilage;
RX MEDLINE=96325055; PubMed=8702629;
RA Hu B., Trinh K., Figueira W.F., Price P.A.;
RT "Isolation and sequence of a novel human chondrocyte protein related
RT to mammalian members of the chitinase protein family.";
RL J. Biol. Chem. 271:19415-19420(1996).
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM (SHOWN HERE) AND A
CC LONG FORM; MAY BE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN CHONDROCYTES, FOLLOWED
CC BY SYNIOCYTES, LUNG AND HEART. NOT DETECTED IN BRAIN, SPLEEN,
CC PANCREAS, AND LIVER.
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U58515; AAB04534.1; -
DR EMBL; U58514; AAB04533.1; -
DR EMBL; U49835; AAC50597.1; ALT_INIT.
DR MIM; 601526; -
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE-NEG.
KW Hydrolase; Signal; Alternative splicing.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 390 CHITINASE 3-LIKE PROTEIN 2.
FT CARBOHYD 35 35 N-LINKED (GLCNAC... ) (POTENTIAL).
FT VARSPPLIC 1 12 MGAT7MDQKSLW -> SQQWLQSGPDEVWLYLVCEHTHIFT
FT FT AITWDSRTREGSDTDFSFHPSI (IN LONG ISOFORM).
SQ SEQUENCE 390 AA; 43500 MW; 97B86A2F3AA35677 CRC64;

Query Match 56.1%; Score 78; DB 1; Length 390;
Best Local Similarity 56.5%; Pred. No. 0.00014;
Matches 13; Conservative 4; Mismatches 6; Indels 0; Gaps 0;

QY 1 YKLCVYTSWSQVREGDGSXFPD 23
Db 27 YKLCVFTNWSQDRCEPKFTPE 49
||||:||||:|
|

RESULT 6
OGP_MOUSE STANDARD; PRT; 721 AA.
ID OGP_MOUSE
AC Q62010;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN).
GN OVGPI OR OGP OR CHIT5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Cviduct;

```

```
RX MEDLINE=96115001; PubMed=7492680;
RA Sendai Y., Komiya H., Suzuki K., Onuma T., Kikuchi M., Hoshi H.,
RA Araki Y.;
RT "Molecular cloning and characterization of a mouse oviduct-specific
RT glycoprotein."
RL Biol. Reprod. 53:285-294(1995).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -!- TISSUE SPECIFICITY: EPITHELIAL CELLS OF THE OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D32137; BAA06863.1; -.
CC MGD; MGI:106661; Chits.
CC InterPro; IPR001579; Chitinase_2.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
CC KW Glycoprotein; Fertilization; Repeat; Signal.
CC FT SIGNAL 1 21 POTENTIAL.
CC FT DOMAIN 486 632 21 X 7 AA TANDEM REPEATS OF S-K-T-T-
CC [TAP]-G-[IV].
CC FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 469 469 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT SEQUENCE 721 AA; 78807 MW; 37246C8F01665652 CRC64;
CC
CC Query Match 54.0%; Score 75; DB 1; Length 721;
CC Best Local Similarity 52.0%; Pred. No. 0.00075;
CC Matches 13; Conservative 3; Mismatches 9; Indels 0; Gaps 0;
CC
CC QY 1 YKLCVYTSWSQYREGDGSXPFDAL 25
CC |||||:|:| | | | |
CC DB 22 YKLCVYFTNWAHSRPGPASTMPHDL 46
CC
CC RESULT 7
CC OGP_PAPAN STANDARD; PRT; 623 AA.
CC AC P36718;
CC DT 01-JUN-1994 (Rel. 29, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 15-JUL-1999 (Rel. 38, Last annotation update)
CC DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
CC DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN).
CC GN OVGP1 OR OGP.
CC OS Papio anubis (Olive baboon).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
CC OC Cercopithecoidea; Papio.
CC OX NCBI_TaxID=9555;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE-Oviduct;
CC RX MEDLINE=98244335; PubMed=9584944;
CC RA Verhage H.G., Fazleabas A.T., Mavrogianis P.A., O'Day-Bowman M.B.,
CC RA Donnelly K.M., Arias E.B., Jaffe R.C.;
CC RT "The baboon oviduct: characteristics of an oestradiol-dependent
CC RT oviduct-specific glycoprotein."
CC RL Hum. Reprod. Update 3:541-552(1997).
CC RN [2]
CC RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
CC RC -TISSUE-Oviduct;
```

```
RX MEDLINE=91367180; PubMed=1716345;
RA Donnelly K.M., Fazleabas A.T., Verhage H.G., Mavrogianis P.A.,
RA Jaffe R.C.;
RT "Cloning of a recombinant complementary DNA to a baboon (Papio
RT anubis) estradiol-dependent oviduct-specific glycoprotein."
RL Mol. Endocrinol. 5:356-364(1991).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- DEVELOPMENTAL STAGE: AT THE TIME OF OVULATION.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M59903; AAB39765.1; -.
CC PIR; A37954; A37954.
CC InterPro; IPR001579; Chitinase_2.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
CC KW Glycoprotein; Fertilization; Signal.
CC FT SIGNAL 1 21 BY SIMILARITY.
CC FT CHAIN 22 623 OVIDUCT-SPECIFIC GLYCOPROTEIN.
CC FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC FT SEQUENCE 623 AA; 69291 MW; 9E2ICE481FFF1268 CRC64;
CC
CC Query Match 50.4%; Score 70; DB 1; Length 623;
CC Best Local Similarity 48.0%; Pred. No. 0.0038;
CC Matches 12; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
CC
CC QY 1 YKLCVYTSWSQYREGDGSXPFDAL 25
CC :|||:|:| | | | |
CC DB 22 HKLCVYFTNWAHSRPGPASILPHDL 46
CC
CC RESULT 8
CC OGP_HUMAN STANDARD; PRT; 678 AA.
CC AC Q12889; Q15841;
CC DT 01-NOV-1997 (Rel. 35, Created)
CC DT 01-NOV-1997 (Rel. 35, Last sequence update)
CC DT 20-AUG-2001 (Rel. 40, Last annotation update)
CC DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
CC DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN).
CC GN OVGP1 OR OGP.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A.
CC RC TISSUE-Oviduct;
CC RX MEDLINE=95119256; PubMed=7819450;
CC RA Arias E.B., Verhage H.G., Jaffe R.C.;
CC RT "Complementary deoxyribonucleic acid cloning and molecular
CC RT characterization of an estrogen-dependent human oviductal
CC RT glycoprotein."
CC RL Biol. Reprod. 51:685-694(1994).
CC RN [2]
CC RP SEQUENCE FROM N.A.
CC RA Jaffe R.C.;
CC RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
CC CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
```

CC -!- SUBCELLULAR LOCATION: SECRETORY GRANULES.
 CC -!- TISSUE SPECIFICITY: OVIDUCT.
 CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U09550; AAB86946.1; -
 DR EMBL: U58010; AAB04126.1; -
 DR EMBL: U58001; AAB04126.1; JOINED.
 DR EMBL: U58002; AAB04126.1; JOINED.
 DR EMBL: U58003; AAB04126.1; JOINED.
 DR EMBL: U58004; AAB04126.1; JOINED.
 DR EMBL: U58005; AAB04126.1; JOINED.
 DR EMBL: U58006; AAB04126.1; JOINED.
 DR EMBL: U58007; AAB04126.1; JOINED.
 DR EMBL: U58008; AAB04126.1; JOINED.
 DR EMBL: U58009; AAB04126.1; JOINED.
 DR MIM: 603578; -
 DR InterPro: IPR001579; Chitinase_2.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR PROSITE: PS01095; CHITINASE_18; FALSE_NEG.
 KW Glycoprotein; Fertilization; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 678 OVIDUCT-SPECIFIC GLYCOPROTEIN.
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 596 596 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 648 648 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 648 648 M -> T (IN REF. 2).
 FT CONFLICT 477 477 S -> P (IN REF. 2).
 FT CONFLICT 511 511 S -> P (IN REF. 2).
 FT CONFLICT 514 514 Y -> H (IN REF. 2).
 FT CONFLICT 676 676 E -> Q (IN REF. 2).
 FT SEQUENCE 678 AA; 75421 MW; 245F2CEDCE92768B CRC64;
 SQ
 Query Match 50.4%; Score 70; DB 1; Length 678;
 Best Local Similarity 48.0%; Pred. No. 0.0041;
 Matches 12; Conservative 4; Mismatches 9; Indels 0; Gaps 0;
 QY 1 YKLVCYTTWSQYREGDGSXFPDL 25
 Db 22 HKLVCFYTNWASRPASLPHDL 46
 :|||||:|:| | | | |
 RESULT 9
 OGP_SHEEP
 ID OGP_SHEEP STANDARD; PRT; 539 AA.
 AC Q28542; Q28543;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
 DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN) (ESTRUS-ASSOCIATED
 DE OVIDUCTAL GLYCOPROTEIN) (OEGP).
 GN OVGPI OR OGP
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-39.
 RC TISSUE=Oviduct;
 RX MEDLINE=95269691; PubMed=7750470;
 RA Desouza M.M., Murray M.K.;

RT "An estrogen-dependent secretory protein, which shares identity with
 RT chitinases, is expressed in a temporally and regionally specific
 RT manner in the sheep oviduct at the time of fertilization and embryo
 RT development.";
 RN Endocrinology 136:2485-2496(1995).
 RL [2]
 RN SEQUENCE OF 10-539 FROM N.A.
 RC STRAIN=MERINO; TISSUE=Oviduct;
 RX MEDLINE=96329120; PubMed=8726871;
 RA Marshall J.T.A., Nancarrow C.D., Brownlee A.G.;
 RT "Cloning and sequencing of a cDNA encoding an ovine
 RT oestrus-associated oviductal protein.";
 RL Reprod. Fertil. Dev. 8:305-310(1996).
 CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
 CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
 CC -!- SUBCELLULAR LOCATION: SECRETORY GRANULES.
 CC -!- TISSUE SPECIFICITY: OVIDUCT.
 CC -!- DEVELOPMENTAL STAGE: LEVELS ARE HIGHEST IN THE FIMBRIA AND AMPULLA
 CC AT ESTRUS AND ON DAY 1 OF PREGNANCY, WHEN GAMETE TRANSPORT AND
 CC FERTILIZATION OCCURS IN THE E2-DOMINATED FALLOPIAN TUBE. LEVELS
 CC DECLINE SIGNIFICANTLY ON DAY 2 AND UNDERGO A FURTHER SIGNIFICANT
 CC REDUCTION ON DAY 3 OF PREGNANCY COINCIDENT WITH TRANSPORT OF THE
 CC EMBRYO FROM THE OVIDUCT TO THE UTERUS. A REPRODUCTIVE STAGE
 CC ASSOCIATED WITH RISING PROGESTERONE LEVELS.
 CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
 CC -----
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: U16719; AAC48471.1; -
 DR EMBL: U17988; AAB01052.1; -
 DR InterPro: IPR001579; Chitinase_2.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR PROSITE: PS01095; CHITINASE_18; FALSE_NEG.
 KW Glycoprotein; Fertilization; Signal.
 FT SIGNAL 1 21
 FT CHAIN 22 539 OVIDUCT-SPECIFIC GLYCOPROTEIN.
 FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CONFLICT 13 13 M -> V (IN REF. 2).
 FT CONFLICT 122 122 K -> N (IN REF. 2).
 FT CONFLICT 282 282 A -> V (IN REF. 2).
 FT CONFLICT 375 375 A -> V (IN REF. 2).
 FT CONFLICT 484 484 R -> H (IN REF. 2).
 FT CONFLICT 520 520 I -> T (IN REF. 2).
 FT SEQUENCE 539 AA; 59535 MW; F35000269987C193 CRC64;
 SQ
 Query Match 48.2%; Score 67; DB 1; Length 539;
 Best Local Similarity 50.0%; Pred. No. 0.0093;
 Matches 11; Conservative 4; Mismatches 7; Indels 0; Gaps 0;
 QY 1 YKLVCYTTWSQYREGDGSXFP 22
 Db 22 HKLVCFYTNWASRPASILP 43
 :|||||:|:| | | | |
 RESULT 10
 CHIT_CAEEL
 ID CHIT_CAEEL STANDARD; PRT; 617 AA.
 AC Q11174;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE PUTATIVE ENDOCHITINASE (EC 3.2.1.14).
 GN C04F6.3.
 OS Caenorhabditis elegans.
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;

RT "Isolation and characterization of a genomic clone for the gene of an
 RT insect molting enzyme, chitinase.";
 RL Insect Biochem. Mol. Biol. 27:37-47(1997).
 CC -!- FUNCTION: DIGEST CHITIN IN THE EXOSKELETON DURING THE MOLTING
 CC PROCESS.
 CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF
 CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
 CC -!- SUBCELLULAR LOCATION: SECRETED.
 CC -!- TISSUE SPECIFICITY: EPIDERMIS AND GUT.
 CC -!- DEVELOPMENTAL STAGE: HIGH LEVELS SEEN IN THE EPIDERMIS ON DAY 0,
 CC BUT RAPIDLY DISAPPEARS AND IS UNDETECTED ON DAYS 1-4 OF FIFTH
 CC INSTAR. IT REAPPEARS ON DAY 5 AND PEAKS ON DAY 7 AFTER WHICH A
 CC RAPID DECLINE IS SEEN. IN THE GUT IT IS DETECTED ON DAY 6 WITH LOWER
 CC LEVELS SEEN ON DAYS 0, 7 AND 8.
 CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U02270; AAC04924.1; -;
 DR EMBL; L49234; AAB53952.1; -;
 DR InterPro; IPR002557; Chitin_binding.
 DR InterPro; IPR001579; Chitinase_2.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF01607; Chitin_bind_2; 1.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR SMART; SM00494; ChtBD2; 1.
 DR PROSITE; PS01095; CHITINASE_18; 1.
 DR Hydrolyase; Glycosidase; Chitin degradation; Signal; Glycoprotein.
 KW SIGNAL
 FT CHAIN 1 19 POTENTIAL.
 FT CHAIN 20 554 ENDOCHITINASE.
 FT DOMAIN 396 453 SER/THR-RICH.
 FT ACT_SITE 146 146 PROTON DONOR (BY SIMILARITY).
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 85 85 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 407 407 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 554 AA; 62203 MW; 3989D756C96CD490 CRC64;

Query Match 42.4%; Score 59; DB 1; Length 554;
 Best Local Similarity 47.1%; Pred. No. 0.16;
 Matches 8; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 2 KLVCYTTSWSQYREGDG 18
 :|||||:|:|
 DB 24 RIVCFNSWAVYREGVG 40

RESULT 15
 MAG9_HUMAN
 ID MAG9_HUMAN STANDARD; PRT; 315 AA.
 AC P43362; Q92910;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE MELANOMA-ASSOCIATED ANTIGEN 9 (MAGE-9 ANTIGEN).
 GN MAGEA9 OR MAGE9.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95012457; PubMed=7927540;
 RA de Plaen E., Arden K., Traversari C., Gaforio J.J., Szikora J.-P.,
 de Smet C., Brasseur F., van der Bruggen P., Lethe B., Lurquin C.,

RA Brasseur R., Chomez P., de Backer O., Cavenee W., Boon T.;
 RT "Structure, chromosomal localization, and expression of 12 genes of
 the MAGE family.";
 RL Immunogenetics 40:360-369(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Timms K.M., Bondeson M.L., Ansari-Lari M.A., Lagerstedt K.,
 RA Nelson D.L., Pettersson U., Gibbs R.A.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: NOT KNOWN, THOUGH MAY PLAY A ROLE IN EMBRYONAL
 CC DEVELOPMENT AND TUMOR TRANSFORMATION OR ASPECTS OF TUMOR
 CC PROGRESSION.
 CC -!- TISSUE SPECIFICITY: EXPRESSED IN MANY TUMORS OF SEVERAL TYPES,
 CC SUCH AS MELANOMA, HEAD AND NECK SQUAMOUS CELL CARCINOMA, LUNG,
 CC CARCINOMA AND BREAST CARCINOMA, BUT NOT IN NORMAL TISSUES EXCEPT
 CC FOR TESTES AND PLACENTA.
 CC -!- SIMILARITY: BELONGS TO THE MAGE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; U10694; AAA68877.1; -;
 DR EMBL; U66083; AAB67888.1; -;
 DR InterPro; IPR002190; MAGE.
 DR Pfam; PF01454; MAGE; 1
 KW Antigen; Multigene family; Tumor antigen.
 FT DOMAIN 34 37 POLY-GLU.
 FT DOMAIN 87 90 POLY-GLU.
 SQ SEQUENCE 315 AA; 35088 MW; 7FD2ED10D680D928 CRC64;

Query Match 38.1%; Score 53; DB 1; Length 315;
 Best Local Similarity 64.3%; Pred. No. 0.73;
 Matches 9; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 6 YTTWSQYREGDGS 19
 ||| |||:|
 DB 72 YTTLSQFDEGSSS 85

Search completed: December 7, 2001, 00:39:46
 Job time: 900 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:23:21 : Search time 135.48 Seconds
(without alignments)
26.991 Million cell updates/sec

Title: US-09-164-862B-1
Perfect score: 139
Sequence: 1 YKLVYVTSWSQYREGDGSXFPDAL 25

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL_17.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phase.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|--------|---------------------|
| 1 | 134 | 96.4 | 383 | Q29411 | Q29411 sus scrofa |
| 2 | 128 | 92.1 | 352 | Q9WTV1 | Q9WTV1 rattus norv |
| 3 | 118 | 84.9 | 332 | Q18949 | Q18949 bos taurus |
| 4 | 116 | 83.5 | 381 | Q99J84 | Q99J84 mus musculus |
| 5 | 90 | 64.7 | 1635 | Q17412 | Q17412 aedes aegypt |
| 6 | 88 | 63.3 | 472 | Q9JLNI | Q9JLNI mus musculus |
| 7 | 88 | 63.3 | 473 | Q9D803 | Q9D803 mus musculus |
| 8 | 88 | 63.3 | 473 | Q99PH2 | Q99PH2 mus musculus |
| 9 | 85 | 61.2 | 504 | Q9NGK8 | Q9NGK8 wuchereria |
| 10 | 84 | 60.4 | 476 | Q9BZP6 | Q9BZP6 homo sapien |
| 11 | 83 | 59.7 | 387 | Q9H3V8 | Q9H3V8 homo sapien |
| 12 | 83 | 59.7 | 466 | Q13231 | Q13231 homo sapien |
| 13 | 83 | 59.7 | 527 | Q91773 | Q91773 penaeus jap |
| 14 | 82 | 59.0 | 43 | Q9UDJ8 | Q9UDJ8 homo sapien |
| 15 | 82 | 59.0 | 458 | Q9W5U2 | Q9W5U2 drosophila |
| 16 | 82 | 59.0 | 497 | Q25615 | Q25615 onchocerca |
| 17 | 82 | 59.0 | 520 | Q17100 | Q17100 acanthocheil |
| 18 | 82 | 59.0 | 524 | Q27919 | Q27919 acanthocheil |
| 19 | 81 | 58.3 | 508 | Q9W5U3 | Q9W5U3 drosophila |

| | | | | | |
|----|------|------|-----|-----------|---------------------|
| 20 | 57.6 | 1080 | 5 | Q9BLI6 | Q9BLI6 bombyx mori |
| 21 | 78 | 696 | 5 | Q9VZV2 | Q9VZV2 drosophila |
| 22 | 73 | 52.5 | 37 | Q9TBJ1 | Q9TBJ1 brugia paha |
| 23 | 70 | 50.4 | 398 | 11 P70201 | P70201 mus musculus |
| 24 | 70 | 50.4 | 398 | 11 Q35744 | Q35744 mus musculus |
| 25 | 70 | 50.4 | 624 | 6 O19118 | O19118 macaca mula |
| 26 | 69 | 49.6 | 226 | 6 Q9XT30 | Q9XT30 macaca radi |
| 27 | 69 | 49.6 | 396 | 11 Q9D701 | Q9D701 mus musculus |
| 28 | 68 | 48.9 | 120 | 11 Q99MJ5 | Q99MJ5 mus musculus |
| 29 | 68 | 48.9 | 483 | 5 Q23737 | Q23737 chelonius sp |
| 30 | 68 | 48.9 | 498 | 5 Q9W223 | Q9W223 drosophila |
| 31 | 65 | 46.8 | 572 | 5 Q26042 | Q26042 penaeus jap |
| 32 | 65 | 46.8 | 620 | 5 Q9Y0D4 | Q9Y0D4 penaeus mon |
| 33 | 63 | 45.3 | 405 | 5 Q97403 | Q97403 phaeton coc |
| 34 | 61 | 43.9 | 514 | 5 P90547 | P90547 entamoeba i |
| 35 | 59 | 42.4 | 467 | 5 O15993 | O15993 penaeus jap |
| 36 | 59 | 42.4 | 543 | 5 Q9GV05 | Q9GV05 bombyx mori |
| 37 | 59 | 42.4 | 543 | 5 Q9GR93 | Q9GR93 bombyx mori |
| 38 | 59 | 42.4 | 544 | 5 Q9GQC4 | Q9GQC4 bombyx mori |
| 39 | 59 | 42.4 | 552 | 5 Q9GV44 | Q9GV44 spodoptera |
| 40 | 59 | 42.4 | 553 | 5 P91731 | P91731 hyphantria |
| 41 | 59 | 42.4 | 565 | 5 P90710 | P90710 bombyx mori |
| 42 | 59 | 42.4 | 565 | 5 Q9GPG9 | Q9GPG9 bombyx mand |
| 43 | 59 | 42.4 | 595 | 5 Q9VFR3 | Q9VFR3 drosophila |
| 44 | 58 | 41.7 | 525 | 5 O44079 | O44079 anopheles g |
| 45 | 57 | 41.0 | 507 | 5 P90546 | P90546 entamoeba h |

ALIGNMENTS

| | | | | | |
|----------|------------------------------------------------------------------------|--------------|------|---------|--|
| RESULT 1 | | | | | |
| Q29411 | | | | | |
| ID | Q29411 | PRELIMINARY; | PRT; | 383 AA. | |
| AC | Q29411; | | | | |
| DT | 01-NOV-1996 (TREMBLrel. 01, Created) | | | | |
| DT | 01-NOV-1996 (TREMBLrel. 01, Last sequence update) | | | | |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last annotation update) | | | | |
| DE | 38 KDA HEPARIN-BINDING GLYCOPROTEIN. | | | | |
| OS | Sus scrofa (Pig). | | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| OC | Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus. | | | | |
| OX | NCBI_TaxID=9823; | | | | |
| RN | [1] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE-SMOOTH MUSCLE: | | | | |
| RX | MEDLINE=95286589; PubMed=7768902; | | | | |
| RA | Shackleton L.M., Mann D.M., Millis A.J.T.; | | | | |
| RT | "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in | | | | |
| RT | differentiating vascular smooth muscle cells as a member of a group of | | | | |
| RT | proteins associated with tissue remodeling." | | | | |
| RL | J. Biol. Chem. 270:13076-13083(1995). | | | | |
| RN | [2] | | | | |
| RP | SEQUENCE FROM N.A. | | | | |
| RC | TISSUE-SMOOTH MUSCLE: | | | | |
| RA | Shackleton L.M., Mann D.M., Millis A.J.T.; | | | | |
| RL | Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases. | | | | |
| DR | EMBL; U19900; AAA86482.1; -. | | | | |
| DR | EMBL; Z47803; CAA87784.1; -. | | | | |
| DR | InterPro; IPR001223; Glyco_hydro_18. | | | | |
| DR | Pfam; PF00704; Glyco_hydro_18; 1. | | | | |
| KW | Heparin-binding. | | | | |
| SQ | SEQUENCE 383 AA; 42443 MW; 3D1039F49910BDC4 CRC64; | | | | |

Query Match 96.4%; Score 134; DB 6; Length 383;

Best Local Similarity 92.0%; Pred. No. 7.9e-13;

Matches 23; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 YKLVYVTSWSQYREGDGSXFPDAL 25

Db 22 YKLVYVTSWSQYREGDGSXFPDAI 46

RESULT 2
 Q9WTV1 ID Q9WTV1 PRELIMINARY; PRT; 352 AA.
 AC Q9WTV1
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE GLYCOPROTEIN-39 PRECURSOR (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LEWIS;
 RA Wendling U., Boots A.M.H., van Eden W.;
 RT "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
 potential autoantigen in arthritis."
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF062038; AAD22610.1; -
 DR HSP; P07254; ICTN.
 DR InterPro; IPR000677; 2S_Globulin.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR PRINTS; PR00551; 2S_GLOBULIN.
 FT NON_TER 1
 FT NON_TER 352
 SQ SEQUENCE 352 AA; 39391 MW; CBDE991610AC936C CRC64;

Query Match 92.1%; Score 128; DB 11; Length 352;
 Best Local Similarity 88.0%; Pred. No. 6.2e-12;
 Matches 22; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 YKLCYITWSQYREGDGSXFPDAL 25
 |||||:|||||:|||||
 Db 1 YKLCYITWSQYREGDGSXFPDAL 25

RESULT 3
 O18949 ID O18949 PRELIMINARY; PRT; 332 AA.
 AC O18949;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE CHITINASE-LIKE PROTEIN 1 (FRAGMENT).
 GN CLP-1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Recklies A.D., White C.;
 RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF011373; AAB64304.1; -
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR NON_TER 1
 FT NON_TER 332
 SQ SEQUENCE 332 AA; 37434 MW; 03F163323486408F CRC64;

Query Match 84.9%; Score 118; DB 6; Length 332;
 Best Local Similarity 90.9%; Pred. No. 2.1e-10;
 Matches 20; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 4 VCYYTWSQYREGDGSXFPDAL 25
 |||||:|||||:|||||
 Db 1 VCYYTWSQYREGDGSXFPDAL 22

RESULT 4
 Q99J84 ID Q99J84 PRELIMINARY; PRT; 381 AA.
 AC Q99J84;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE SIMILAR TO CHITINASE 3-LIKE 1 (CARTILAGE GLYCOPROTEIN-39).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC005611; AAH05611.1; -
 DR EMBL; BC003780; AAH03780.1; -
 DR EMBL; BC004734; AAH04734.1; -
 SQ SEQUENCE 381 AA; 42979 MW; EF6588C5AE9D4450 CRC64;

Query Match 83.5%; Score 116; DB 11; Length 381;
 Best Local Similarity 80.0%; Pred. No. 5e-10;
 Matches 20; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 YKLCYITWSQYREGDGSXFPDAL 25
 |||||:|||||:|||||
 Db 22 YKLCYITWSQYREGDGSXFPDAL 46

RESULT 5
 O17412 ID O17412 PRELIMINARY; PRT; 1635 AA.
 AC O17412;
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE PROBABLE CHITINASE 2 (EC 3.2.1.14).
 GN CHT2.
 OS Aedes aegypti (Yellowfever mosquito).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;
 OC Aedes
 OX NCBI_TaxID=7159;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=98324849; PubMed=9662472;
 de la Vega H., Specht C.A., Liu Y., Robbins P.W.;
 RT "Chitinases are a multi-gene family in Aedes, Anopheles and
 Drosophila."
 RL Insect Mol. Biol. 7:233-239(1998).
 CC -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF
 CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).
 DR EMBL; AF026492; AAB81850.1; -
 DR InterPro; IPR001579; Chitinase_2.
 DR InterPro; IPR002557; Chitin_binding.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF01607; Chitin_bind_2; 3.
 DR Pfam; PF00704; Glyco_hydro_18; 3.
 DR SMART; SM00494; ChitBD2; 3.

| | |
|-------------------------------------------------------------|-----------------------------------------------------------------------------------|
| DR | PROSITE; PS01095; CHITINASE_18; 3. |
| KW | Hydrolase; Glycosidase; Chitin degradation; Glycoprotein; |
| KW | Multigene family. |
| FT | CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD 749 749 N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD 890 890 N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD 1338 1338 N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD 1479 1479 N-LINKED (GLCNAC. . .) (POTENTIAL). |
| SQ | SEQUENCE 1635 AA; 185993 MW; EAL16F83AAC129FA CRC64; |
| Query Match 64.7%; Score 90; DB 5; Length 1635; | |
| Best Local Similarity 56.5%; Pred. No. 2.7e-05; | |
| Matches 13; Conservative 6; Mismatches 4; Indels 0; Gaps 0; | |
| Qy | 1 YKLCVYTSWSOYREGDSXFPD 23 |
| Dd | 1257 FKVCYFTNNAWYRQGDKYTPD 1279 |
| : : : | |
| RESULT 6 | |
| Q9JLNI | PRELIMINARY; PRT; 472 AA. |
| ID | Q9JLNI AC Q9JLNI; |
| DT | 01-OCT-2000 (Tremblrel. 15, Created) |
| DT | 01-OCT-2000 (Tremblrel. 15, Last sequence update) |
| DT | 01-JUN-2001 (Tremblrel. 17, Last annotation update) |
| DE | PUTATIVE CHITINASE PRECURSOR (FRAGMENT). |
| GN | 2200003E03RIK OR YNL. |
| OS | Mus musculus (Mouse). |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus. |
| OX | NCBI_TaxID=10090; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN-AFCCL476668; TISSUE-SKIN; |
| RA | Price P.A., Harris S.C., Williamson M.K.; |
| RL | *YNL, A Putative Mouse Chitinase.*; |
| RT | Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases. |
| CC | -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL HYDROLASES). |
| CC | EMBL; AF154571; AAP31644.1; -. |
| DR | MGD; MG1:1913535; 2200003E03RIK. |
| DR | InterPro; IPR001579; Chitinase.2. |
| DR | InterPro; IPR002557; Chitin_binding. |
| DR | InterPro; IPR001223; Glyco_hydro_18. |
| DR | Pfam; PF01607; Chitin_bind_2; 1.. |
| DR | Pfam; PF00704; Glyco_Hydro_18; 1.. |
| DR | SMART; SM00494; ChtBD2; 1. |
| DR | PROSITE; PS01095; CHITINASE_18; 1. |
| KW | Glycosidase; Hydrolase; Signal. |
| FT | NON_TER 1 1 |
| FT | SIGNAL <1 20 POTENTIAL. |
| FT | CHAIN 21 472 PUTATIVE CHITINASE. |
| SQ | SEQUENCE 472 AA; 51872 MW; FF59088512C8A7F0 CRC64; |
| Query Match 63.3%; Score 88; DB 11; Length 472; | |
| Best Local Similarity 60.9%; Pred. No. 1.4e-05; | |
| Matches 14; Conservative 4; Mismatches 5; Indels 0; Gaps 0; | |
| Qy | 1 YKLCVYTSWSOYREGDSXFPD 23 |
| Dd | 21 YNLICFTNNAQYRPLGSKFPD 43 |
| : : : | |
| RESULT 7 | |
| Q9D803 | PRELIMINARY; PRT; 473 AA. |
| ID | Q9D803 AC Q9D803; |
| DT | 01-JUN-2001 (Tremblrel. 17, Created) |

RA Bijl N., Moe C., Place A., Aerts J.M.F.G.;
 RT Identification of a Novel Acidic Mammalian Chitinase Distinct from
 RL Chitriose.
 RT Chitriose.
 RL J. Biol. Chem. 276:6770-6778 (2001).
 RL EMBL; AF290003; AAG60018.1; -
 KW Hydrolase; Glycosidase.
 SO SEQUENCE 473 AA; 51977 MW; 389D87557BEC0784 CRC64;

| | | | | |
|-----------------------|----------------------------|---------------|-----------|-------------|
| Query Match | 63.3% | Score 88; | DB 11; | Length 473; |
| Best Local Similarity | 60.9%; | Pred. No. | 1.4e-05; | |
| Matches 14; | Conservative 4; | Mismatches 5; | Indels 0; | Gaps 0; |
| Qy | 1 YKLVCYVTSWSQYREGDGSXPFD | 23 | | |
| | : : : | | | |
| bh | 22 YNLICYFTNWAQYRPLGLGSPFD | 44 | | |

```

RESULT      9
Q9NGK8      PRELIMINARY;      PRRT;      504 AA.
ID
Q9NGK8;
AC      01-OCT-2000 (TrEMBLrel. 15, Created)
DT      01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT      01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE      CURRICULAR ENDOCHITINASE.
OS      Wuchereria bancrofti.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Spirurida; Filarioidea;
OC      Onchocercidae; Wuchereria.
OX      NCBI_TaxID=6293;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Varatharajulu V., Kubofcik J., Williams S.A., Kaliraj P.,
RA      Jayaraman K., Numan T.;
RT      "Molecular characterization of endochitinase from Wuchereria bancrofti
RT      microfilarial cDNA library.";
RL      Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF250997; AAF66988.1; -.
DR      InterPro; IPRO01579; Chitinase_2.
DR      InterPro; IPRO02557; Chitin_binding.
DR      InterPro; IPRO01223; Glyco_hydro_18.
DR      Pfam; PF01607; Chitin_bind_2; 1.
DR      Pfam; PF00704; Glyco_hydro_18; 1.
DR      SMART; SM00494; ChtBD2; 1.
DR      PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
SO      SEQUENCE      504 AA; 56480 MW; D65C888300DE2C27 CRC64;

```

| Query Match | Score 85; | DB 5; | Length 504; |
|-----------------------|--------------------------|-----------------|---------------------------------|
| Best Local Similarity | Pred. No. 4.5e-05; | | |
| Matches | 12; | Conservative 4; | Mismatches 3; Indels 0; Gaps 0; |
| QY 5 | CYVTSWSQYREGDGSXFPD 23 | | |
| | : : | | |
| pb 27 | CYVTTNNAOYREGGEGKELPE 45 | | |

| | |
|--------|-------------------------------------------------------------|
| RESULT | 10 |
| Q9BZP6 | |
| ID | Q9BZP6 PRELIMINARY; PRT; 476 AA. |
| AC | Q9BZP6; |
| DT | 01-JUN-2001 (TREMBlrel. 17, Created) |
| DT | 01-JUN-2001 (TREMBlrel. 17, Last sequence update) |
| DT | 01-JUN-2001 (TREMBlrel. 17, Last annotation update) |
| DE | ACIDIC MAMMALIAN CHITINASE PRECURSOR (EC 3.2.1.14). |
| OS | Homo sapiens (Human) |
| OC | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; |
| CC | Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. |
| OX | NCBI_TaxID=9606; |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE-STOMACH, AND LUNG; |
| RX | PubMed=11085997; |

RA Boot R.G., Blommaert E.F.C., Swart E., Chauharali-van Der Vlugt K.,
RA Bijl N., Moe C., Place A., Aerts J.M.F.G.;
"Identification of a Novel Acidic Mammalian Chitinase Distinct from
Chitinase 1".
RA Chitinase 1.
RL J. Biol. Chem. 276:6770-6778 (2001).
RL EMBL; AF290004; AAG60019.1; -.
DR Signal; Hydrolase; Glycosidase.
KW SIGNAL 1 21 POTENTIAL.
FT SIGNAL 476 AA; 52271 MW; 92527BAD2F7E54CC CRC64;
FT SO SEQUENCE

| | | | | |
|-----------------------|--------------|-----------|------------|-------------|
| Query Match | 60.4%; | Score 84; | DB 4; | Length 476; |
| Best Local Similarity | 56.5%; | Pred. No. | 6e-05; | |
| Matches 13; | Conservative | 4; | Mismatches | 6; |
| | | | Indels | 0; |
| | | | Gaps | 0; |

```

QY      1 YKLCYYTWSQYREGDGSXFPD 23
        | : | | : | : | | | | |
Db     22 YOLTCYFTNWAQYRPGIGREMPD 44

```

| RESULT | 11 |
|--------|-------------------------------------------------------------------|
| Q9H3V8 | |
| ID | Q9H3V8 |
| | PRELIMINARY; PRT; 387 AA. |
| AC | Q9H3V8; |
| DT | 01-MAR-2001 (TREMBLrel. 16, Created) |
| DT | 01-MAR-2001 (TREMBLrel. 16, Last sequence update) |
| DT | 01-JUN-2001 (TREMBLrel. 17, Last annotation update) |
| DE | CHITOTRIOSIDASE PRECURSOR. |
| OS | Homo sapiens (Human); |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| OX | NCBI_TaxID=9606; |

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96064695; PubMed=7592832;
 RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J., Aerts J.M.;
 RT "Cloning of a CDNA encoding chitotriosidase, a human chitinase
 RL J. Biol. Chem. 270:26252-26256(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98421482; PubMed=9748235;
 RA Boot R.G., Renkema G.H., Verhoeck M., Strijland A., Blik J.,
 RA De Meulemeester T.M., Mannens M.M., Aerts J.M.;
 RT "The human chitotriosidase gene. Nature of inherited enzyme
 RL J. Biol. Chem. 273:25680-25685(1998).
 CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).

| CC | HYDROLASES} | EMBL; U62662; AAG10644.1; - | INTERPRO; IPR001579; Chitinase_2. | DR | InterPro; IPR001223; Glyco_hydro_18. | DR | Pfam; PF00704; Glyco_hydro_18; 1. | DR | PROSITE; PS01095; CHITINASE_18; 1. | DR | Glycosidase; Hydrolase; Signal. | KW | POTENTIAL. | FT | SIGNAL | FT | CHAIN | SEQUENCE | 367 AA; 43133 MW; 03A272B8BC5E0D71 CRC64; | SO | CHITOTRIOSIDASE. |
|----|-------------|-----------------------------|-----------------------------------|----|--------------------------------------|----|-----------------------------------|----|------------------------------------|----|---------------------------------|----|------------|----|--------|----|-------|----------|-------------------------------------------|----|------------------|
| CC | HYDROLASES} | EMBL; U62662; AAG10644.1; - | INTERPRO; IPR001579; Chitinase_2. | DR | InterPro; IPR001223; Glyco_hydro_18. | DR | Pfam; PF00704; Glyco_hydro_18; 1. | DR | PROSITE; PS01095; CHITINASE_18; 1. | DR | Glycosidase; Hydrolase; Signal. | KW | POTENTIAL. | FT | SIGNAL | FT | CHAIN | SEQUENCE | 367 AA; 43133 MW; 03A272B8BC5E0D71 CRC64; | SO | CHITOTRIOSIDASE. |

| | | | | |
|-----------------------|--------------|--------------------|---------------|-------------|
| Query Match | 59.7% | Score 83; | DB 4; | Length 387; |
| Best Local Similarity | 54.2% | Pred. No. 6.9e-05; | | |
| Matches 13; | Conservative | 5; | Mismatches 6; | Indels 0; |
| | | | | Gaps 0; |

| | | | |
|----|----|-------------------------|----|
| Qy | 2 | KLVCYYTWSQYREGDGSXFPDAL | 25 |
| | | : : : : : | |
| pb | 23 | KLVCYFTNWAQYRGEARELPKDL | 46 |

RESULT 12
Q13231
ID Q13231
AC Q13231;
PRELIMINARY;
PRT; 466 AA.

```

01-NOV-1996 (TrEMBLrel. 01, Created)
01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
CHITOTRIOSIDASE PRECURSOR.
DE Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=960646595; PubMed=7592832;
RA Boot R.G., Renkema G.H., Strijland A., van Zonneveld A.J.,
RA Aarts J.M.F.G.;
RT Cloning of a cDNA encoding chitotriosidase, a human chitinase
RT produced by macrophages.";
RL J. Biol. Chem. 270:26252-26256(1995).
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCO
CC HYDROLASES).
CC EMBL: U29615; AAC50246.1; -.
DR InterPro: IPR001579; Chitinase_2.
DR InterPro: IPR002557; Chitin_binding.
DR InterPro: IPR001223; Glyco_hydro_18.
DR Pfam: PF01607; Chitin_bind_2; 1.
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR SMART: SM00494; ChtBD2; 1.
DR PROSITE: PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase; Signal.
FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 466
SQ SEQUENCE 466 AA; 51681 MW; B4312D1C888E386D CRC64;

```

Query Match 59.7%; Score 83; DB 4; Length 466;
Best Local Similarity 54.2%; Pred. NO. 8.5e-05;
Matches 13; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

| | | | |
|----|----|--------------------------|----|
| Qy | 2 | KLVCYYTWSQYREGDGSXFPDAL | 25 |
| | | : : : : : : | |
| Db | 23 | KLVCYFTNWAQYRQGEARFLPKDL | 46 |

```

RESULT 13
P91773
ID P91773 PRELIMINARY; PRT; 527 AA.
AC P91773;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE PCHI-2.
OS Penaeus japonicus (Kuruma prawn).
OC Eukaryota; Metazoa; Arthropoda; Crustacea; Malacostraca;
OC Eumalacostraca; Eucarida; Decapoda; Dendrobranchiata; Penaeoidea;
OC Penaeidae; Penaeus.
OC NCBI_TaxId=27405;
RN [1]
RP SEQUENCE FROM N. A.
RA Watanabe T., Kono M.;
RT "Isolation of a cDNA Encoding a Chitinase Family Protein from
RT Cuticular tissues of the Kuruma Prawn Penaeus japonicus.";
RL Zool. Sci. 0:0-0(1996).
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCO
CC HYDROLASES).
DR EMBL: D89751; BAA14014.1; -.
DR InterPro: IPR001579; Chitinase_2.
DR InterPro: IPR002557; Chitin_binding.
DR InterPro: IPR001223; Glyco_hydro_18.
DR Pfam: PF00704; Glyco_hydro_18; 1.
DR SMART: SM00494; Chtd2; 1.
DR PROSITE: PS01095; CHITINASE_18; 1.
DR Glycosidase; Hydrolase.
KW SEQUENCE 527 AA; 59162 MW; B9CBAEAB8CDBF8710 CRC64;
SQ

```

| | | | | |
|-----------------------|-------------------------------------------------------------|-------------------------|----------|-------------------------|
| Query Match | 59.7%; | Score 83; | DB 5; | Length 527; |
| Best Local Similarity | 52.2%; | Pred. No. 9.7e-05; | | |
| Matches | 12; | Conservative | 6; | Mismatches 5; |
| | | | | Indels |
| QY | 1 | YKLVCVYTSMSOYREGDGSXFPD | 23 | |
| | | : : : : : : : : | | |
| Db | 151 | YKVVCFYFNWIRGSGKYRPE | 173 | |
| | | : : : : : : : | | |
| RESULT | 14 | | | |
| Q9UDJ8 | | | | |
| ID | Q9UDJ8 | PRELIMINARY; | PRT; | 43 AA. |
| AC | Q9UDJ8; | | | |
| DT | 01-MAY-2000 (TEMBLrel. 13, Created) | | | |
| DT | 01-MAY-2000 (TEMBLrel. 13, Last sequence update) | | | |
| DT | 01-JUN-2001 (TEMBLrel. 17, Last annotation update) | | | |
| DE | CHITOTRIOSIDASE (FRAGMENTS). | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eutele- | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo; | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE. | | | |
| RX | MEDLINE=95138187; PubMed=7836450; | | | |
| RA | Renkema G.H., Boot R.G., Mulijers A.O., Donker-Koopman W.E. | | | |
| RA | Aerts J.M.; | | | |
| RT | "Purification and characterization of human chitotriosidase | | | |
| RT | member of the chitinase family of proteins."; | | | |
| RL | J. Biol. Chem. 270:2198-2202(1995). | | | |
| DR | InterPro: IPR001223; Glyco_hydro_18. | | | |
| DR | Pfam: PF00704; Glyco_hydro_18; 1. | | | |
| ET | NON_TER | 1 | | |
| ET | NON_TER | 43 | | |
| SQ | SEQUENCE | 43 AA; | 4796 MW; | 34D58D34A6513BE1 CRC64; |

Query Match 59.0%; Score 82; DB 4; Length 43;
Best Local Similarity 57.1%; Pred. No. 9.2e-06;
Matches 12; Conservative 5; Mismatches 4; Indels

| | | | |
|--------|-----------------------------------------------------------------------|-----------------------------------------|---------|
| Qy | 2 | KLVCYNTSWSOYREGDSXRP | 22 |
| | | : : : : | |
| Dd | 2 | KLVCFYTNWAOYRQGEARFLP | 22 |
| | | | |
| RESULT | 15 | | |
| Q9W5U2 | | PRELIMINARY; | |
| ID | Q9W5U2 | PRT; | 458 AA. |
| AC | Q9W5U2 | OL7422; | |
| DT | 01-MAY-2000 | (TREMBLrel. 13, Created) | |
| DD | 01-MAY-2000 | (TREMBLrel. 13, Last sequence update) | |
| DT | 01-JUN-2001 | (TREMBLrel. 17, Last annotation update) | |
| DE | PROBABLE CHITINASE 3 (EC 3.2.1.14). | | |
| GN | CMT3 OR CG18140. | | |
| OS | Drosophila melanogaster (Fruit fly). | | |
| OC | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; | | |
| OC | Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; | | |
| OX | Ephydroidea; Drosophilidae; Drosophilla. | | |
| OX | NCBI_TaxID=7227; | | |
| RN | [1] | | |
| RP | SEQUENCE FROM N.A. | | |
| RC | STRAIN=BERKELEY; | | |
| RX | MEDLINE=20196006; PubMed=10731132; | | |
| RA | Adams W.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D., | | |
| RA | Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F., | | |
| RA | George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N., | | |
| RA | Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X., | | |
| RA | Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D., | | |
| RA | Van K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G., | | |
| RA | Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D., | | |
| RA | Ballew R.M., Basu A., Bonaldo J., Bayraktaroglu L., Beasley E.M., | | |
| RA | Beeson K.Y., Benos P.V., Berham B.P., Bhandari D., Bolshakov S., | | |
| RA | Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P., | | |

Tue Dec 11 08:46:44 2001

us-09-164-862b-1.rspt

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacheb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE OF 182-294 FROM N.A.
RN STRAIN-CANTON-S;
RC MEDLINE-98324849; PubMed=9662472;
RX de la Vega H., Specht C.A., Liu Y., Robbins P.W.;
RA "Chitinases are a multi-gene family in *Aedes*, *Anopheles* and
RT *Drosophila*.";
RT Insect Mol. Biol. 7:233-239(1998).
RL -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF N-
CC ACTYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC -1- HYDROLASES).
CC EMBL; AE002743; AAF45395.1; -.
DR EMBL; AF026502; AAB81860.1; -.
DR FlyBase; FBgn0022701; Cht3.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR002557; Chitin_binding.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF01607; Chitin_bind_2; 2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR SMART; SM00494; ChtBD2; 2.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Hydrolase; Glycosidase; Chitin degradation; Glycoprotein;
KW Multigene family.
KW ACT-SITE 295 295 PROTON DONOR (BY SIMILARITY).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 254 254 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 395 395 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 428 428 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 458 AA; 52330 MW; 4A063190B7E96248 CRC64;

Query Match 59.0%; Score 82; DB 5; Length 458;
Best Local Similarity 52.2%; Pred. No. 0.00012;
Matches 12; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

QY 1 YKLVYVTSQVREGDSXFPD 23
||::||:|:|:|:|:|:|
Db 173 YKVICYFTNNAWYRKIGIGRETPD 195

Search completed: December 7, 2001, 00:38:25
Job time: 904 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:23:14 ; Search time 113.83 seconds
(without alignments)
12.364 Million cell updates/sec

Title: US-09-164-862b-2

Perfect score: 96

Sequence: 1 LNTLKNRPNLKTLLSVGG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 522463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_1101.*
1: /SIDSB/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | ID | Description |
|------------|-------|-------------|--------|---------------------------|
| 1 | 96 | 100.0 | 19 16 | AA1980 internal pe |
| 2 | 96 | 100.0 | 19 16 | AA1981 internal pe |
| 3 | 96 | 100.0 | 19 21 | AA1982 Human cancer marke |
| 4 | 96 | 100.0 | 126 21 | AA1983 Human secreted pro |
| 5 | 96 | 100.0 | 321 9 | AA1984 Polypeptide involv |
| 6 | 96 | 100.0 | 383 18 | AA1985 Bovine whey protei |
| 7 | 96 | 100.0 | 383 21 | AA1986 Gp38k protein sequ |
| 8 | 80 | 83.3 | 16 18 | AA1987 Human cartilage gl |
| 9 | 69 | 71.9 | 399 22 | AA1988 F. venenatum endoc |
| 10 | 68 | 70.8 | 108 21 | AA1989 Human secreted pro |
| 11 | 68 | 70.8 | 385 19 | AA1990 Amino acid sequenc |

| | | | | | | |
|----|----|------|-----|----|----------|--------------------|
| 12 | 68 | 70.8 | 385 | 19 | AAW47033 | Human cartilage gp |
| 13 | 68 | 70.8 | 416 | 19 | AAW47034 | Human cartilage gp |
| 14 | 68 | 70.8 | 421 | 19 | AAW53121 | Amino acid sequenc |
| 15 | 68 | 70.8 | 421 | 19 | AAW53122 | Amino acid sequenc |
| 16 | 68 | 70.8 | 423 | 18 | AAW35930 | Human cartilage gl |
| 17 | 64 | 66.7 | 357 | 21 | AAW47948 | Arabidopsis thalia |
| 18 | 64 | 66.7 | 379 | 21 | AAW47947 | Arabidopsis thalia |
| 19 | 64 | 66.7 | 383 | 21 | AAW47946 | Arabidopsis thalia |
| 20 | 64 | 66.7 | 410 | 21 | AAW47939 | Arabidopsis thalia |
| 21 | 64 | 66.7 | 432 | 21 | AAW47938 | Arabidopsis thalia |
| 22 | 64 | 66.7 | 455 | 22 | AAW62541 | Disease treatment |
| 23 | 64 | 66.7 | 476 | 22 | AAW62543 | Disease treatment |
| 24 | 64 | 66.7 | 537 | 16 | AAW73991 | Bovine oviduct spe |
| 25 | 64 | 66.7 | 718 | 16 | AAW73992 | Murine oviduct spe |
| 26 | 63 | 65.6 | 16 | 18 | AAW45109 | Human cartilage gl |
| 27 | 63 | 65.6 | 373 | 19 | AAW40262 | Human chitinase pr |
| 28 | 63 | 65.6 | 373 | 19 | AAW40261 | Human chitinase pr |
| 29 | 63 | 65.6 | 373 | 20 | AAW42427 | Clone of the C-ter |
| 30 | 63 | 65.6 | 373 | 20 | AAW42428 | Chitinase amino ac |
| 31 | 63 | 65.6 | 377 | 22 | AAW62545 | Disease treatment |
| 32 | 63 | 65.6 | 387 | 18 | AAW08585 | Human 39 kDa chiti |
| 33 | 63 | 65.6 | 398 | 22 | AAW62544 | Disease treatment |
| 34 | 63 | 65.6 | 464 | 18 | AAW31497 | Human chitotriosid |
| 35 | 63 | 65.6 | 466 | 18 | AAW31498 | Human chitotriosid |
| 36 | 63 | 65.6 | 466 | 18 | AAW08584 | Human 50 kDa chiti |
| 37 | 63 | 65.6 | 466 | 19 | AAW40259 | Human chitinase pr |
| 38 | 63 | 65.6 | 466 | 19 | AAW40260 | Human chitinase pr |
| 39 | 63 | 65.6 | 466 | 20 | AAW42425 | MO-218 clone of hu |
| 40 | 63 | 65.6 | 466 | 20 | AAW42426 | MO-13B clone of hu |
| 41 | 63 | 65.6 | 466 | 22 | AAE00432 | Human chitinase pr |
| 42 | 63 | 65.6 | 466 | 22 | AAE00433 | Human chitinase pr |
| 43 | 63 | 65.6 | 668 | 16 | AAW73993 | Hamster oviduct sp |
| 44 | 62 | 64.6 | 451 | 21 | AAW20843 | Arabidopsis thalia |
| 45 | 62 | 64.6 | 455 | 21 | AAW20842 | Arabidopsis thalia |

ALIGNMENTS

RESULT 1
AAW70752
ID AAW70752 standard; peptide; 19 AA.
XX
AC AAW70752;
XX
DT 27-SEP-1995 (first entry)
XX
DE YKL-40 internal peptide A.
XX
KW YKL-40; internal peptide A; connective tissue degradation;
KW osteoporosis; tumour metastasis; polysaccharide hydrolase;
KW chitinase; joint disease; competitive immunoassay.
XX
OS Homo sapiens.
XX
PN WO9502188-A.
XX
PD 19-JAN-1995.
XX
PF 12-JUL-1993; 93WO-US06579.
XX
PR 12-JUL-1993; 93WO-US06579.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Johansen JS, Price PA;
XX
DR WPI; 1995-066992/09.
XX
PT Assay for YKL-40 as a marker for the degradation of connective
PT tissue - is a competitive immunoassay, and can be used to
PT identify joint disease and the extent of tumour metastasis

PS Disclosure; Page 34; 5lpp; English.

XX AAR70751, AAR70752 and AAR70753 are the N-terminal, and internal
CC peptides A and B respectively of the human osteosarcoma cell
CC line MG63 derived protein YKL-40. These peptides show a sequence
CC homology with a bacterial polysaccharide hydrolase (chitinase),
CC this suggests that the protein YKL-40 is involved in connective
CC tissue degradation (CTD). By performing a competitive immunoassay
CC for YKL-40 (as a marker of CTD) joint disease, osteoporosis and
CC the extent of tumour metastasis can be diagnosed.

XX Sequence 19 AA;

Query Match 100.0%; Score 96; DB 16; Length 19;

Best Local Similarity 100.0%; Pred. No. 3.4e-09; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

QY 1 LNTLKNRPNLKTLLSVGG 19
Db 1 Intlnrnpnlktllsvgg 19

RESULT 2

AAR70746
ID AAR70746 standard; Peptide; 19 AA.

AC AAR70746;

XX 12-AUG-1995 (first entry)

XX YKL-40 internal peptide A.

XX YKL-40; diagnosis; prognosis; therapy; breast cancer; metastasis;
KW marker; joint disease; connective tissue.

XX Homo sapiens.

XX WO9501995-A.

XX 19-JAN-1995.

XX 08-JUL-1994; 94WO-US07754.

XX 09-JUL-1993; 93US-0089989.

XX (REGC) UNIV CALIFORNIA.

XX Johansen JS, Price PA;

XX WPI; 1995-066866/09.

XX Use of YKL-40 and anti-YKL-40 antibodies - for developing prods.
PT for diagnosis prognosis and therapy of diseases involving
PT connective tissue degradation.

XX Disclosure; Page 64; 88pp; English.

XX YKL-40 (40 kDa) was purified from human osteosarcoma MG63 cells.
CC The N-terminal sequence is shown in AAR70745; the full coding region
CC of the YKL-40 gene is given in AAQ85245. Homology of the N-terminal
CC and 2 internal peptides (AAR70746-47) with a bacterial polysaccharide
CC hydrolase suggests that YKL-40 degrades polysaccharide components
CC of connective tissue. YKL-40 is a marker of e.g. metastatic breast
CC cancer and inflammatory or degenerative joint diseases.

XX Sequence 19 AA;

Query Match 100.0%; Score 96; DB 16; Length 19;

Best Local Similarity 100.0%; Pred. No. 3.4e-09; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;

QY 1 LNTLKNRPNLKTLLSVGG 19
Db 1 Intlnrnpnlktllsvgg 19

RESULT 3

AAY79492
ID AAY79492 standard; Peptide; 19 AA.

XX AAY79492;

XX 01-AUG-2000 (first entry)

XX Human cancer marker YKL-40 internal peptide A.

XX YKL-40; human; lung cancer; bronchus cancer; colorectal cancer;
KW prostate cancer; breast cancer; pancreas cancer; stomach cancer;
KW ovary cancer; bladder cancer; brain cancer; oesophagus cancer;
KW cervix cancer; melanoma; uterine endometrial cancer;
KW oral cavity cancer; pharynx cancer; liver cancer; kidney cancer;
KW biliary tract cancer; small bowel cancer; appendix cancer;
KW salivary gland cancer; thyroid gland cancer; testis cancer;
KW adrenal gland cancer; osteosarcoma; chondrosarcoma; liposarcoma;
KW malignant fibrous histiocytoma; infection; pneumonia; meningitis;
KW arthritis; rheumatoid arthritis; osteoarthritis; fibrosis;
KW liver cirrhosis; marker; diagnosis; prognosis.

XX Homo sapiens.

XX WO200019206-A1.

XX 06-APR-2000.

XX 29-SEP-1999; 99WO-US22615.

XX 01-OCT-1998; 98US-0164862.

XX (REGC) UNIV CALIFORNIA.

XX Price PA, Johansen JS;

XX WPI; 2000-303485/26.

XX Novel methods for detecting cancers and evaluating the prognosis of
XX cancer using YKL-40 as a marker of cancer

XX Disclosure; Page 101; 11lpp; English.

XX This peptide represents internal peptide A of human YKL-40
XX mature polypeptide. YKL-40 is a 40 kDa protein having Tyr, Lys and
XX Leu as its N-terminal residues (hence, YKL-40). The protein can be
XX obtained from osteosarcoma cell line MG63. The present sequence for
XX correlates to 14/19 residues of an internal amino acid sequence for
XX chitinase. Thus, YKL-40 is a mammalian member of the chitinase
XX family. It is suggested that YKL-40 degrades the polysaccharide
XX components in connective tissue and/or is a lectin that binds to
XX specific glycan structures in the extracellular environment of
XX cells. YKL-40 is useful as a marker for the presence or absence of
XX a cancer and for the prognosis of a cancer. A claimed method for
XX estimating survival length of cancer patients comprises obtaining a
XX biological sample from the cancer patient and measuring the level
XX of YKL-40 in the sample, a higher level than in healthy humans
XX being indicative of reduced survival expectancy. The biological
XX sample is obtained from a cancer patient having at least a
XX preliminary diagnosis of cancer selected from lung, bronchus,
XX colorectal, prostate, breast, pancreas, stomach, ovary, urinary
XX bladder, brain, central nervous system, peripheral nervous system,
XX oesophagus, cervix, melanoma, uterine endometrial, oral cavity,
XX pharynx, liver, kidney, biliary tract, small bowel, appendix,
XX salivary gland, thyroid gland, testes, or adrenal gland cancer, or
XX osteosarcoma, chondrosarcoma, liposarcoma, or malignant fibrous
XX histiocytoma. Levels of the YKL-40 marker are elevated in
XX pathologies associated with tissue remodeling, e.g. degenerative

CC bone diseases such as rheumatoid arthritis, osteoarthritis, fibrosis,
CC cirrhosis of the liver, and cancer, especially breast, colon, high
CC prostate, or lung cancer. The marker can be used to identify high
CC risk patients, and so allow selection of appropriate therapeutic
CC regimens. The methods may also be used to detect bacterial
CC infections, such as bacterial pneumonia and meningitis, as these
CC cause an elevation in YKL-40 levels, as well as diseases
CC characterized by macrophage activation, e.g. giant cell arteritis.
CC The YKL-40 marker may also be used to evaluate treatment efficacy,
CC to check for recurrence of a cancer, to monitor terminal phase
CC patients, and to check the efficacy of surgical removal of a
CC primary tumor. The methods allow estimation of the survival time
CC of patients with cancers, especially prostate, lung or colorectal
CC cancer, where the colorectal cancer is Duke's stage A, B, C, or D.
XX
SQ Sequence 19 AA;

Query Match 100.0%; Score 96; DB 21; Length 19;
Best Local Similarity 100.0%; Pred. No. 3.4e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LNTLKNRPNLKTLLSVGG 19
| | | | | | | | | | | | | | | | | | |
Db 1 Intlknrpnlktllsvgg 19

RESULT 4
AAG00227
ID AAG00227 standard; Protein; 126 AA.
XX
AC AAG00227;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 4308.
XX
KW Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EP1033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PA (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
XX WPI; 2000-500381/45.
DR N-PSDB; AAC00233.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX Claim 13; SEQ ID 4308; 7lpp + CD-ROM; English.
PS
XX

The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic

CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
SQ Sequence 126 AA;

Query Match 100.0%; Score 96; DB 21; Length 126;
Best Local Similarity 100.0%; Pred. No. 3.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LNTLKNRPNLKTLLSVGG 19
| | | | | | | | | | | | | | | | | | |
Db 80 Intlknrpnlktllsvgg 98

RESULT 5
AAP81342
ID AAP81342 standard; Protein; 321 AA.
XX
AC AAP81342;
XX
DT 19-OCT-1990 (first entry)
XX
DE Polypeptide involved in protective mechanisms.
XX
KW Immune response; cell growth.
XX
PN JP63032898-A.
XX
PD 30-JAN-1988.
XX
PF 16-JUL-1986; 86JP-0167518.
XX
PR 16-JUL-1986; 86JP-0167518.
XX
PA (DAIN) DAINIPPON PHARM KK.
XX
XX WPI; 1988-068419/10.
DR P-PSDB; AAP81342.
XX
XX New polypeptide and DNA encoding it - related to protective
PT mechanisms such as immune response etc.
PT
PS Disclosure; ; P; Japanese.
XX
XX This polypeptide is involved in protective mechanisms such as immune
CC response, cell growth and activation of protective functions.
CC
XX
SQ Sequence 321 AA;

Query Match 100.0%; Score 96; DB 9; Length 321;
Best Local Similarity 100.0%; Pred. No. 9.1e-08;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 LNTLKNRPNLKTLLSVGG 19
| | | | | | | | | | | | | | | | | | |
Db 79 Intlknrpnlktllsvgg 97

RESULT 6
AAW26751
ID AAW26751 standard; Protein; 383 AA.
XX
AC AAW26751;
XX
DT 11-MAY-1998 (first entry)
XX
DE Bovine whey protein.
XX
KW Bovine whey protein; human cartilage glycoprotein 39; HC gp-39;
KW autoantigen; antigen; autoimmune disease; rheumatoid arthritis;

Tue Dec 11 08:46:45 2001

KW Inflammation; arthritogenic protein; immunotherapy; therapy.

XX Bos taurus.

XX WO9740149-A1.

XX 30-OCT-1997.

XX 15-APR-1997; 97WO-EP01903.

XX 18-APR-1996; 96US-0634493.

XX (ALKU) AKZO NOBEL NV.

XX Boots AMH, Bos ES, Verheijden GFW;

XX WPI; 1997-535833/49.

XX N-PSDB; AAT99452.

XX Autoantigen proteins homologous to human cartilage glycoprotein 39 -
PT induce arthritis and provide antigen-specific treatment of articular
PT cartilage destruction in autoimmune diseases e.g. rheumatoid
PT arthritis

XX Claim 11; Page 23-25; 35pp; English.

XX This bovine 39 kDa whey protein sequence was deduced from cDNA
CC clones (see AAT99452) isolated from a cattle mammary gland cDNA
CC library using human cartilage glycoprotein 39 (HC gp-39) cDNA as
CC probe. Bovine whey protein was shown to be arthritogenic, inducing
CC arthritis in animals in the same way as described for HC gp-39 (see
CC AAW26750). Arthritogenic proteins such as bovine whey protein, and
CC which display at least 50% amino acid homology to HC gp-39, are
CC very suitable for inducing systemic tolerance of the immune system
CC to homologous autoantigens and can be used to delay and/or suppress
CC arthritic development in mammals. They can induce specific T-cell
CC tolerance to HC gp-39 in patients suffering from T-cell mediated
CC cartilage destruction, such as rheumatoid arthritis. Arthritogenic
CC proteins are also suitable to induce arthritis in animals,
CC preferably mice, e.g. for use in drug screening.

XX Sequence 383 AA;

Query Match 100.0%; Score 96; DB 18; Length 383;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTLKNRPNLKTLSSVGG 19
Db 80 Lntiknrpnkltlsvgg 98

RESULT 7
AAB03442
ID AAB03442 standard; Protein; 383 AA.

XX AAB03442;

XX 03-JAN-2001 (first entry)

XX Gp38k protein sequence.

XX Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW cancer; vascular trauma; vascular disease; atherosclerosis; restenosis;

XX Unidentified.

XX Key Location/Qualifiers
XX Peptide 1..24

XX /label= signal_peptide

XX 30...37

XX /note= "peptide antibody"

FT

FT Modified-site 59..61
FT /label= glycosylation_site
FT Binding-site 68..75
FT /label= leucine_zipper
FT Binding-site 82..89
FT /label= leucine_zipper
FT Active-site 131..136
FT /label= chitinase_active_site
FT Binding-site 143..146
FT /label= heparin_binding_site
FT Binding-site 147..154
FT /label= hyaluronic_acid_binding_site
FT Binding-site 262..270
FT /label= hyaluronic_acid_binding_site
FT Binding-site 278..281
FT /label= glycosaminoglycan_binding_site
FT Region 354..357
FT /note= "acidic region"
FT Binding-site 368..376
FT /label= hyaluronic_acid_binding_site

XX WO200034469-A1.

XX 15-JUN-2000.

XX 10-DEC-1999; 99WO-US29262.

XX 11-DEC-1998; 98US-0111856.

XX (UUNY) UNIV NEW YORK STATE RES FOUND.

XX Millis AJT;

XX WPI; 2000-431300/37.

XX Clusterin and gp38k-related peptide capable of altering cell migration
PT useful for treating atherosclerosis, cancer and stenosis following
PT vascular trauma or disease

XX Disclosure; Fig 2; 43pp; English.

XX The present sequence is the protein sequence of gp38k. Gp38k, a
CC chemoattractant, is essential for the migration of vascular smooth muscle
CC cells (VSMC). The gene and protein can, therefore, be used to promote
CC wound healing, angiogenesis and vasculogenesis, in the treatment of
CC stenosis following vascular trauma or disease and to treat
CC atherosclerosis, and antisense sequences can be used to treat cancer, as
CC angiogenesis is vital for tumour survival.

XX Sequence 383 AA;

Query Match 100.0%; Score 96; DB 21; Length 383;
Best Local Similarity 100.0%; Pred. No. 1.1e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTLKNRPNLKTLSSVGG 19
Db 80 Lntiknrpnkltlsvgg 98

RESULT 8
AAW45110
ID AAW45110 standard; peptide; 16 AA.

XX AAW45110;

XX 28-APR-1998 (first entry)

XX Human cartilage glycoprotein 39 derived peptide #11.

XX Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW immunological tolerance; T-cell; human cartilage glycoprotein 39;

KW HC gp-39; rheumatoid arthritis; epitope.

XX Synthetic.

OS Homo sapiens.

XX WO9740068-A1.

PN 30-OCT-1997.

XX 22-APR-1997; 97WO-EP02051.

PF 24-APR-1996; 96EP-0201106.

XX (ALKU) AKZO NOBEL NV.

PA Boots AMH, Verheijden GFM;

PI WPI; 1997-535775/49.

XX Peptide suitable for use in antigen specific immunosuppressive

PT therapy - resembles or mimics epitope present on HC gp-39, so

PT inducing systemic immunological tolerance to rheumatoid arthritis

PT auto:antigen

XX Disclosure: Page 15; 82pp; English.

XX The present sequence represents a peptide which resembles or mimics an

CC epitope present on human cartilage glycoprotein 39 (HC gp-39), an

CC autoantigen in rheumatoid arthritis. The invention relates to peptides

CC consisting of 16-55 amino acid residues comprising at least one of the

CC following 19 sequences: LVRYTWS; FLCTHIYS; ILYSFANIS; LKTLSSVGG;

CC FIKSVPPFL; FDGLDLAWL; LYPGRDRQ; YDIAKISQ; LDFISIMTY; FLSIMTYDF;

CC FRGEDASP; YAVGYMLRL; MLRLGAPAS; LAYEICDF; LGATVHRT; YLKDRQLAG;

CC LAGAWNAL; VWALDLDF; or LLDLDFQS. They can be used medically in

CC antigen specific immunosuppressive therapy, particularly the treatment

CC of T-cell mediated destruction of articular cartilage in autoimmune

CC diseases (e.g. rheumatoid arthritis). They can also be used to detect

CC activated autoreactive T cells in an individual. The peptides have a

CC specific effect on the autoreactive T cells, thus leaving the other

CC components of the immune system intact, unlike the non-specific

CC suppressive effect of immunosuppressive drugs, and do not cause toxic

CC side effects. The peptides are predominantly recognised by autoreactive

CC T cells from rheumatoid arthritis patients, but rarely by those from

CC healthy donors.

XX Sequence 16 AA;

SQ

Query Match 83.3%; Score 80; DB 18; Length 16;

Best Local Similarity 100.0%; Pred. No. 1.2e-06;

Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TLKRNPNLKTLLSVG 18

Db 1 tlknrpnlnkllsvg 16

RESULT 9

AAU00962

ID AAU00962 standard; Protein; 399 AA.

XX AAU00962;

XX 12-SEP-2001 (first entry)

DE F. venenatum endochitinase polypeptide.

XX Glucanase; endochitinase; exochitinase; cell-wall degradation; fungus;

KW transgenic plant; plant pathogen; bacteria; seafood waste; shell;

KW chitin; chemical modification; glucan.

XX Fusarium venenatum.

OS

XX

PN WO200116353-A1.

XX 08-MAR-2001.

XX 30-AUG-2000; 2000WO-US23802.

XX 30-AUG-1999; 99US-0151582.

PR 11-AUG-2000; 2000US-0224946.

PR 28-AUG-2000; 2000US-0649747.

XX (NOVO) NOVO NORDISK BIOTECH INC.

PA (USDA) US SEC OF AGRIC.

XX Okubara PA, Blechl AE, Hohn TM, Berka RM;

PI WPI; 2001-218524/22.

DR N-PSDB; AAS00796; AAS00800.

XX Fusarium nucleic acids encoding polypeptides having glucanase,

PT endochitinase or exochitinase activity, useful for producing transgenic

PT plants which are resistant to plant pathogens, particularly Fusarium

PT species -

XX Claim 13; Page 167-168; 216pp; English.

XX The sequence represents the Fusarium fungal enzyme, endochitinase.

CC Glucanase, endochitinase and exochitinase are polypeptides with cell-wall

CC degrading activity, derived from Fusarium fungal genes. The associated

CC nucleic acids can be used to produce transgenic plants which are

CC resistant to plant pathogens, particularly Fusarium species. They can

CC also be used to isolate homologous genes from fungi to obtain genes which

CC protect host cells, including fungi, bacteria and plants against related

CC fungal pathogens. The polypeptides, especially chitinases and glucanases,

CC are useful for degrading seafood waste, such as shells that contain

CC chitin, or for chemical modification of chitin or glucan.

XX Sequence 399 AA;

SQ

Query Match 71.9%; Score 69; DB 22; Length 399;

Best Local Similarity 73.7%; Pred. No. 0.0032;

Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLLSVGG 19

Db 83 lnlkrrnlnkllsvg 101

RESULT 10

AAG00129

ID AAG00129 standard; Protein; 108 AA.

XX AAG00129;

XX 06-OCT-2000 (first entry)

DE Human secreted protein, SEQ ID NO: 4210.

XX Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;

KW gene therapy; chromosome mapping.

KW Homo sapiens.

OS EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

PF 26-FEB-1999; 99US-0122487.

XX (GEST) GENSET.

XX

Tue Dec 11 08:46:45 2001

us-09-164-862b-2.rag

PI Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.
DR N-PSDB; AAC00135.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 13; SEQ ID 4210; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.

XX Sequence 108 AA;

Query Match 70.8%; Score 68; DB 21; Length 108;
Best Local Similarity 63.2%; Pred. No. 0.001;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNTLKNRNPMLKTLISVGG 19
Db 85 inslktknpkllisig 103

RESULT 11

ID AAW53123 standard; Protein; 385 AA.

XX AAW53123;

XX 22-JUL-1998 (first entry)

XX Amino acid sequence of a human chitinase alpha-2 protein.

XX Human; chitinase alpha; hpmf51 cDNA clone; antifungal agent; antagonist;
KW antibody; treatment; tissue remodelling disorder; rheumatoid arthritis;
KW atherosclerosis; inflammation; chromosome identification.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..20

FT Protein /note= "signal peptide"

FT Protein 21..385

FT Protein /note= "mature protein"

XX W09806859-AL.

XX 19-FEB-1998.

XX 09-AUG-1996; 96WO-US13003.

XX 09-AUG-1996; 96WO-US13003.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Choi GH, Chopra A, Genzt RL, Rosen CA, Thotakura NR;

XX WPI; 1998-159549/14.

DR N-PSDB; AAV21689.

XX Isolated nucleic acid encoding human chitinase alpha and alpha2 -
PT and related vectors, transformants, proteins and antibodies, useful
PT as antifungal agents and as modulators of chitinase activity for
PT treating tissue re-modelling diseases

XX Claim 2; Fig 5A-B; 96pp; English.

XX The present sequence represents a shortened human chitinase alpha
CC proteinvariant, designated chitinase alpha-2. The nucleotide sequence
CC is obtained by sequencing the hpmf51 cDNA clone. The chitinase protein
CC is potentially useful as an antifungal agent. Antagonists of the protein,
CC e.g. antibodies, are used to treat tissue remodelling disorders,
CC e.g. rheumatoid arthritis, atherosclerosis or inflammation caused by
CC elevated plasma levels of the protein. Nucleic acids encoding the
CC chitinase protein can be used for chromosome identification. Diseases
CC related to abnormal levels of the chitinase protein may be diagnosed
CC by measuring chitinase gene expression in, e.g. serum or urine.

XX Sequence 385 AA;

Query Match 70.8%; Score 68; DB 19; Length 385;
Best Local Similarity 63.2%; Pred. No. 0.0044;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNTLKNRNPMLKTLISVGG 19
Db 80 inslktknpkllisig 98

RESULT 12

AAW47033

ID AAW47033 standard; Protein; 385 AA.

XX AAW47033;

XX 20-JUL-1998 (first entry)

XX Human cartilage gp39-like protein.

XX Human cartilage glycoprotein 39-like gene; HC gp39-L; chitinase;
KW agonist; antagonist; rheumatoid arthritis; osteoarthritis;
KW osteoporosis; atherosclerosis; cancer; metastasis; periodontia;
KW chronic renal disease; therapy; diagnosis.

XX Homo sapiens.

XX EP823478-A2.

XX 11-FEB-1998.

XX 08-AUG-1997; 97EP-0306053.

XX 09-AUG-1996; 96US-0694915.

XX (SMIK) SMITHKLINE BEECHAM CORP.

XX Adamou J, Kirkpatrick RB, Rosenberg M;

XX WPI; 1998-112267/11.

DR N-PSDB; AAV13925.

XX Human cartilage gp39-like protein(s) - used for the treatment of
PT rheumatoid and osteoarthritis, osteoporosis, atherosclerosis,
PT metastatic cancers, periodontia, chronic renal diseases, etc

XX Claim 12; Page 26-28; 41pp; English.

XX This 385-amino acid protein sequence comprises the human
CC cartilage glycoprotein 39-like protein (HC gp39-L) that is a new
CC lymphocyte-associated protein of the chitinase-like family and
CC which is believed to be involved in tissue remodelling. A


```

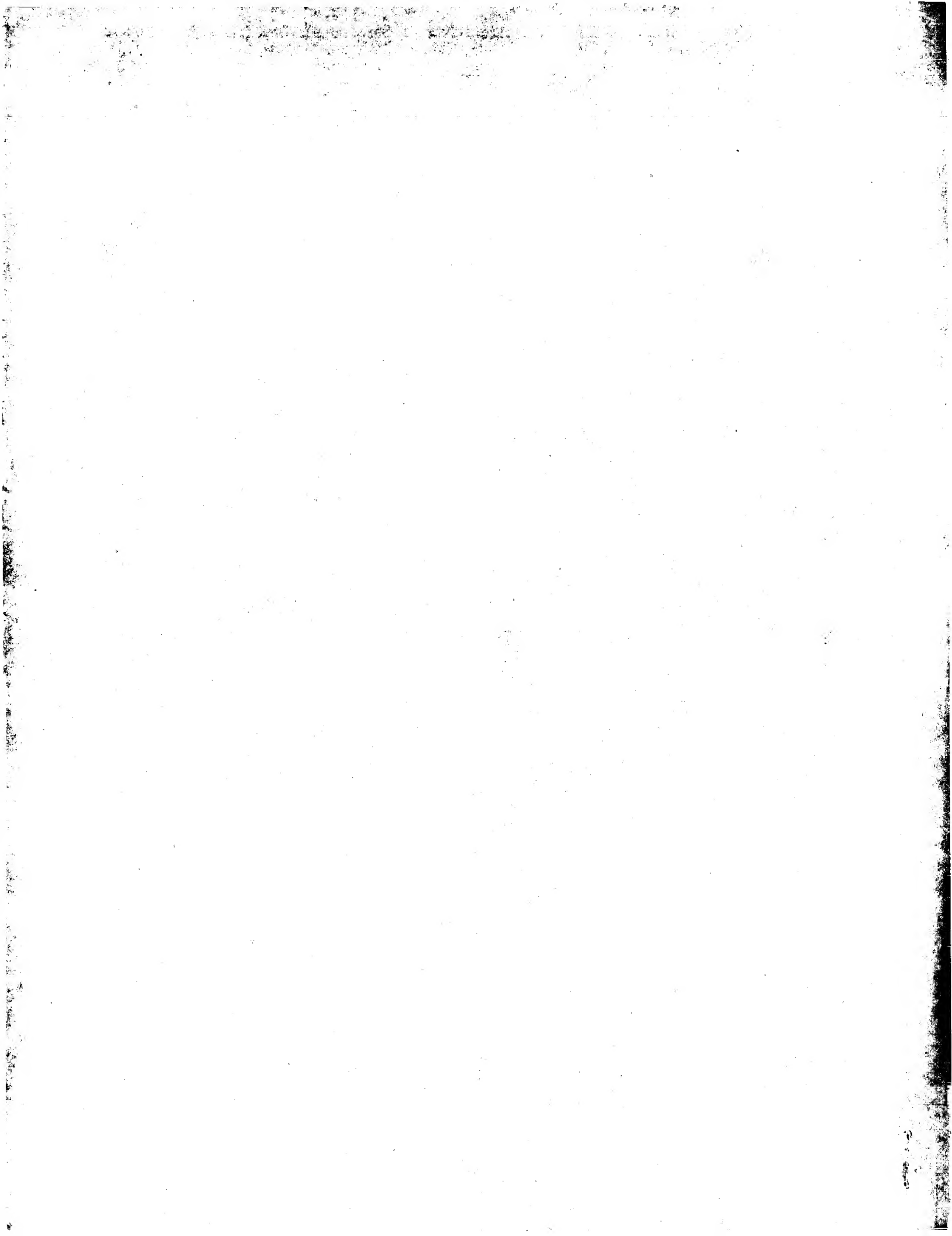
XX      Sequence 421 AA;
SQ
Query Match      70.8%; Score 68; DB 19; Length 421;
Best Local Similarity 63.2%; Pred. No. 0.0049;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY      1 LNTLKNRNPENKTLISVGG 19
      :::: :||| :||| :||| :|||
Db      116 inslktknpkllisig 134
      :||| :||| :||| :||| :|||

Search completed: December 7, 2001, 00:23:15
Job time: 2960 sec

```

| | |
|-----------|--------------------------------------------------------------------------|
| RESULT 15 | |
| AAW53122 | |
| ID | AAW53122 standard; Protein; 421 AA. |
| XX | |
| AC | AAW53122; |
| XX | |
| DT | 22-JUL-1998 (first entry) |
| XX | |
| DE | Amino acid sequence of a human chitinase alpha protein variant. |
| XX | |
| KW | Human; chitinase alpha; hpmf51 cDNA clone; antifungal agent; antagonist; |
| KW | antibody; treatment; tissue remodelling disorder; rheumatoid arthritis; |
| KW | atherosclerosis; inflammation; chromosome identification. |
| XX | |
| OS | Homo sapiens. |
| XX | |
| XX | Key |
| XX | Location/Qualifiers |
| FT | Peptide |
| FT | 1..23 |
| FT | /note= "signal peptide" |
| FT | 24..421 |
| FT | /note= "mature protein" |
| XX | |
| PN | W09806859-A1. |
| XX | |
| PD | 19-FEB-1998. |
| XX | |
| XX | 09-AUG-1996; 96WO-US13003. |
| PF | |
| XX | 09-AUG-1996; 96WO-US13003. |
| PR | |
| XX | (HUMA-) HUMAN GENOME SCI INC. |
| PA | |
| XX | |
| PI | Choi GH, Chopra A, Genzt RL, Rosen CA, Thotakura NR; |
| XX | |
| DR | WPI; 1998-159549/14. |
| DR | N-PSDB; AAV21688. |
| XX | |
| PT | Isolated nucleic acid encoding human chitinase alpha and alpha2 - |
| PT | and related vectors, transformants, proteins and antibodies, useful |
| PT | as antifungal agents and as modulators of chitinase activity for |
| PT | treating tissue re-modelling diseases |
| XX | |
| PS | Claim 2; Fig 2A-B; 96pp; English. |
| XX | |
| CC | The present sequence represents a human chitinase alpha protein variant. |
| CC | The encoded protein has a Ile176 to Glu176 change, compared to the |
| CC | chitinase alpha protein (AAW53121). The nucleotide sequence is obtained |
| CC | by sequencing the hpmf51 cDNA clone. The chitinase protein is |
| CC | potentially useful as an antifungal agent. Antagonists of the protein, |
| CC | e.g. antibodies, are used to treat tissue remodelling disorders, |
| CC | e.g. rheumatoid arthritis, atherosclerosis or inflammation caused by |
| CC | elevated plasma levels of the protein. Nucleic acids encoding the |
| CC | chitinase protein can be used for chromosome identification. Diseases |
| CC | related to abnormal levels of the chitinase protein may be diagnosed |
| CC | by measuring chitinase gene expression in, e.g. serum or urine. |
| XX | |
| SO | Sequence 421 AA; |
| XX | |



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:24:38 ; Search time 71.83 Seconds
(without alignments)
5.952 Million cell updates/sec

Title: US-09-164-862B-2

Perfect score: 96

Sequence: 1 LNTLKNRPNTKLLSVGG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents-AA:*

- 1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep.*
- 2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep.*
- 3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep.*
- 4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep.*
- 5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep.*
- 6: /cgn2_6/ptodata/2/1aa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------------|
| 1 | 96 | 100.0 | 19 | 2 | US-08-581-527-2 |
| 2 | 96 | 100.0 | 19 | 5 | PCT-US94-07754-2 |
| 3 | 80 | 83.3 | 16 | 4 | US-09-171-705-11 |
| 4 | 68 | 70.8 | 385 | 2 | US-08-694-915-2 |
| 5 | 68 | 70.8 | 416 | 2 | US-08-694-915-4 |
| 6 | 63 | 65.6 | 16 | 4 | US-09-171-705-10 |
| 7 | 63 | 65.6 | 373 | 4 | US-09-039-198A-14 |
| 8 | 63 | 65.6 | 373 | 4 | US-09-039-198A-15 |
| 9 | 63 | 65.6 | 387 | 2 | US-08-486-839-6 |
| 10 | 63 | 65.6 | 387 | 3 | US-09-151-011-6 |
| 11 | 63 | 65.6 | 387 | 4 | US-09-343-623-6 |
| 12 | 63 | 65.6 | 466 | 2 | US-08-486-839-4 |
| 13 | 63 | 65.6 | 466 | 3 | US-09-151-011-4 |
| 14 | 63 | 65.6 | 466 | 4 | US-09-039-198A-2 |
| 15 | 63 | 65.6 | 466 | 4 | US-09-039-198A-4 |
| 16 | 63 | 65.6 | 466 | 4 | US-09-343-623-4 |
| 17 | 58 | 60.4 | 866 | 1 | US-08-386-727-8 |
| 18 | 58 | 60.4 | 866 | 2 | US-08-600-452A-8 |
| 19 | 56 | 58.3 | 440 | 3 | US-09-052-778-13 |
| 20 | 55 | 57.3 | 16 | 4 | US-09-171-705-12 |
| 21 | 55 | 57.3 | 452 | 3 | US-09-052-778-15 |
| 22 | 54 | 56.2 | 441 | 3 | US-09-052-778-14 |
| 23 | 53 | 55.2 | 389 | 1 | US-07-939-501A-1 |
| 24 | 53 | 55.2 | 389 | 4 | US-08-448-398-7 |
| 25 | 53 | 55.2 | 423 | 1 | US-07-939-501A-10 |
| 26 | 53 | 55.2 | 423 | 1 | US-07-939-501A-12 |
| 27 | 52 | 54.2 | 371 | 2 | US-08-591-629-2 |

| | | | | | | |
|----|------|------|-----|---|------------------|--------------------|
| 28 | 52 | 54.2 | 377 | 2 | US-08-591-629-8 | Sequence 8, Appl1 |
| 29 | 52 | 54.2 | 424 | 1 | US-08-045-269C-2 | Sequence 2, Appl1 |
| 30 | 52 | 54.2 | 424 | 3 | US-08-371-680-2 | Sequence 2, Appl1 |
| 31 | 52 | 54.2 | 424 | 5 | PCT-US94-01198-2 | Sequence 2, Appl1 |
| 32 | 51 | 53.1 | 442 | 3 | US-09-052-778-2 | Sequence 2, Appl1 |
| 33 | 50 | 52.1 | 428 | 3 | US-09-052-778-12 | Sequence 12, Appl1 |
| 34 | 47 | 49.0 | 561 | 1 | US-08-358-901-2 | Sequence 2, Appl1 |
| 35 | 47 | 49.0 | 561 | 1 | US-08-566-347-2 | Sequence 2, Appl1 |
| 36 | 47 | 49.0 | 561 | 1 | US-08-693-835-2 | Sequence 2, Appl1 |
| 37 | 45 | 46.9 | 44 | 4 | US-09-046-894-41 | Sequence 41, Appl1 |
| 38 | 45 | 46.9 | 44 | 4 | US-09-046-894-42 | Sequence 42, Appl1 |
| 39 | 44 | 45.8 | 348 | 4 | US-09-198-955A-8 | Sequence 8, Appl1 |
| 40 | 42 | 43.8 | 9 | 4 | US-09-171-705-63 | Sequence 63, Appl1 |
| 41 | 42 | 43.8 | 376 | 3 | US-09-025-691-3 | Sequence 3, Appl1 |
| 42 | 42 | 43.8 | 724 | 3 | US-09-121-964-1 | Sequence 1, Appl1 |
| 43 | 41 | 42.7 | 365 | 2 | US-08-204-288-7 | Sequence 7, Appl1 |
| 44 | 40.5 | 42.2 | 363 | 3 | US-08-912-560-2 | Sequence 2, Appl1 |
| 45 | 40 | 41.7 | 454 | 2 | US-09-014-969-9 | Sequence 9, Appl1 |

ALIGNMENTS

RESULT 1

US-08-581-527-2
; Sequence 2, Application US/08581527
; Patent No. 5935798
; GENERAL INFORMATION:
; APPLICANT: Price, Paul A.
; APPLICANT: Johansen, Julia S.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER
; TITLE OF INVENTION: FOR DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/581,527
; FILING DATE: 17-APR-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO95/01995
; FILING DATE: 19-JAN-1995
; APPLICATION NUMBER: 08/089,989
; FILING DATE: 09-JUL-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Freeman, John W.
; REGISTRATION NUMBER: 29,066
; REFERENCE/DOCKET NUMBER: 07341/011001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-542-5070
; TELEFAX: 617-542-8906
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: YKL-40 Internal Peptide A
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1...19
; US-08-581-527-2

us-09-164-862b-2.rai

Tue Dec 11 08:46:45 2001

```
Query Match      100.0%; Score 96; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTLKNRNPMLKTLTLLSVGG 19
   |||||
Db 1 LNTLKNRNPMLKTLTLLSVGG 19

RESULT 2
PCT-US94-07754-2
; Sequence 2, Application PC/TUS9407754
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07754
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD 3665
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: YKL-40 INTERNAL PEPTIDE A
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..19
; PCT-US94-07754-2

Query Match      100.0%; Score 96; DB 5; Length 19;
Best Local Similarity 100.0%; Pred. No. 1.5e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTLKNRNPMLKTLTLLSVGG 19
   |||||
Db 1 LNTLKNRNPMLKTLTLLSVGG 19

RESULT 3
US-09-171-705-11
; Sequence 11, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
; APPLICANT: BOOTS, ANNA M.H.
; APPLICANT: VERHEIJDEN, GILBERTUS F.M.

; TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
; TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
; FILE REFERENCE: O/96198 US
; CURRENT APPLICATION NUMBER: US/09/171,705
; CURRENT FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
; OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
; US-09-171-705-11

Query Match      83.3%; Score 80; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 5.1e-07;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 TLKNRNPMLKTLTLLSVG 18
   |||||
Db 1 TLKNRNPMLKTLTLLSVG 16

RESULT 4
US-08-694-915-2
; Sequence 2, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/694,915
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 34,344
; REFERENCE/DOCKET NUMBER: ATG50017
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
```

ORIGINAL SOURCE:
US-08-694-915-2

Query Match 70.8%; Score 68; DB 2; Length 385;
Best Local Similarity 63.2%; Pred. No. 0.0018;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSVGG 19
DB 80 INSLTKNPKLKILLSIG 98

RESULT 5

US-08-694-915-4
; Sequence 4, Application US/08694915
; Patent No. 5811535
; GENERAL INFORMATION:
; APPLICANT: Adamou, Julie
; APPLICANT: Kirkpatrick, Robert
; APPLICANT: Rosenberg, Martin
; TITLE OF INVENTION: HUMAN CARTILAGE GP39-LIKE GENE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-2799

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 1.5
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/694,915
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Han, William T
REGISTRATION NUMBER: 34,344
REFERENCE/DOCKET NUMBER: ATG50017
TELECOMMUNICATION INFORMATION:
TELEPHONE: 610-270-5219
TELEFAX: 610-270-5090
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 416 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
US-08-694-915-4

Query Match 70.8%; Score 68; DB 2; Length 416;
Best Local Similarity 63.2%; Pred. No. 0.002;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSVGG 19
DB 111 INSLTKNPKLKILLSIG 129

RESULT 6

US-09-171-705-10
; Sequence 10, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
; APPLICANT: BOOTS, ANNA M.H.
; APPLICANT: VERHEIJDEN, GILBERTUS F.M.
; TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
; TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
; FILE REFERENCE: O/96198 US
; CURRENT APPLICATION NUMBER: US/09/171,705
; CURRENT FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
; OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-10

Query Match 65.6%; Score 63; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.0003;
Matches 12; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLK 12
DB 5 LNTLKNRPNLK 16

RESULT 7

US-09-039-198A-14
; Sequence 14, Application US/09039198A
; Patent No. 6200951
; GENERAL INFORMATION:
; APPLICANT: Gray, Patrick W.
; APPLICANT: Tjoelker, Larry W.
; TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 233 South Wacker Drive/6300 Sears Tower
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/039,198A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 27866/34391
TELECOMMUNICATION INFORMATION:
TELEPHONE: (312) 474-6300
TELEFAX: (312) 474-0448
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 373 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-039-198A-14

TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 387 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-151-011-6

Query Match 65.6%; Score 63; DB 3; Length 387;
Best Local Similarity 66.7%; Pred. No. 0.012;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Qy 2 NTLKRNPNLKTLLSVGG 19
 | | | | | | | : | |
Db 81 NGLKKMNPKLKTLLAIGG 98

```

RESULT 11
US-09-343-623-6
; Sequence 6, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:

```

Query Match 65.6%; Score 63; DB 4; Length 387;
Best Local Similarity 66.7%; Pred. NO. 0.012;
Matches 12; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTLKRNPNLKTLLSVGG 19
 | | | | | : | |
Db 81 NGLKKMNPKLKTLAIGG 98

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

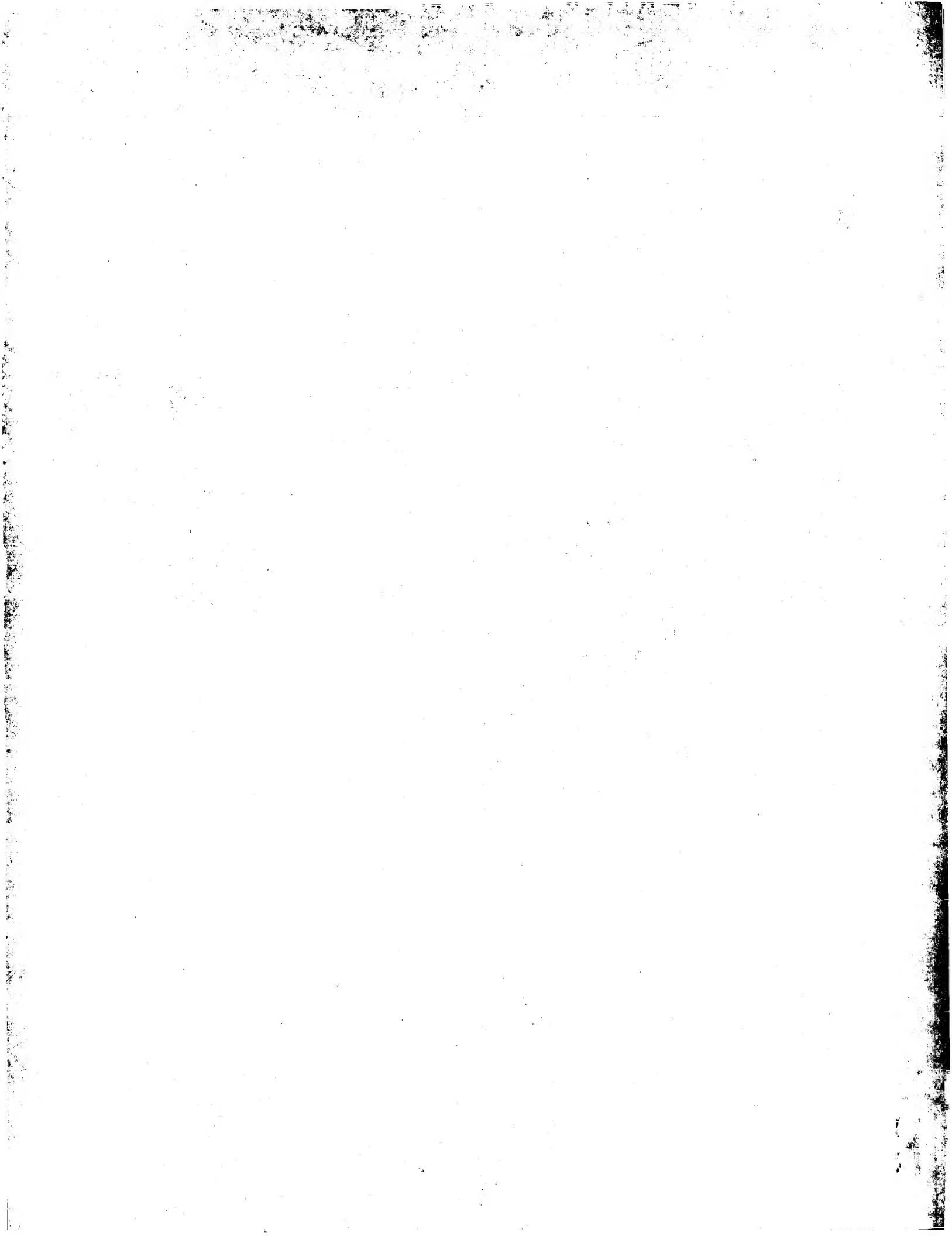
```

| | | | | |
|--------------------------|--------|------------------|-----------|-------------|
| Query Match | 65.6% | Score 63; | DB 2; | Length 466; |
| Best Local Similarity | 66.7%; | Pred. No. 0.015; | | |
| Matches 12: Conservative | 2; | Mismatches 4; | Indels 0; | Gaps 0; |

Qy 2 NTLKRNPNLKTLLSVGG 19
| | | | | : |
Db 81 NGLKKMNPKLKTLAIGG 98

RESULT 13
US-09-151-011-4
: Sequence 4, Application US/09151011
: Patent No. 6057142
: GENERAL INFORMATION:
: APPLICANT: A Human Chitinase, Its Recombinant
: TITLE OF INVENTION: Production, Its Use For Decomposing Chitin, Its Use in
: TITLE OF INVENTION: Therapy or Prophylaxis Against Infection Diseases.
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Hoffmann & Baron, LLP
: STREET: 6900 Jericho Turnpike
: CITY: Syosset
: STATE: New York

:
:
: ADDRESS: Hoffmann & Baron, LLP
:
: STREET: 6900 Jericho Turnpike
:
: CITY: Syosset
:
: STATE: New York
:



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:32:57 ; Search time 401.11 seconds
(without alignments)
13.152 Million cell updates/sec

Title: US-09-164-862B-2
Perfect score: 96
Sequence: 1 LNTLKNRPNLKTLLSVGG 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main : *
1: /cgn2_6/ptodata/1/paa/PTCUS_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US081_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US082_COMB.pep.*
7: /cgn2_6/ptodata/1/paa/US083_COMB.pep.*
8: /cgn2_6/ptodata/1/paa/US084_COMB.pep.*
9: /cgn2_6/ptodata/1/paa/US085_COMB.pep.*
10: /cgn2_6/ptodata/1/paa/US086_COMB.pep.*
11: /cgn2_6/ptodata/1/paa/US087_COMB.pep.*
12: /cgn2_6/ptodata/1/paa/US088_COMB.pep.*
13: /cgn2_6/ptodata/1/paa/US089_COMB.pep.*
14: /cgn2_6/ptodata/1/paa/US090_COMB.pep.*
15: /cgn2_6/ptodata/1/paa/US091_COMB.pep.*
16: /cgn2_6/ptodata/1/paa/US092_COMB.pep.*
17: /cgn2_6/ptodata/1/paa/US093_COMB.pep.*
18: /cgn2_6/ptodata/1/paa/US094_COMB.pep.*
19: /cgn2_6/ptodata/1/paa/US095_COMB.pep.*
20: /cgn2_6/ptodata/1/paa/US096_COMB.pep.*
21: /cgn2_6/ptodata/1/paa/US097_COMB.pep.*
22: /cgn2_6/ptodata/1/paa/US098_COMB.pep.*
23: /cgn2_6/ptodata/1/paa/US099_COMB.pep.*
24: /cgn2_6/ptodata/1/paa/US60_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|----------------------|
| 1 | 96 | 100.0 | 19 | 4 | US-08-089-989-2 |
| 2 | 96 | 100.0 | 19 | 15 | US-09-164-862B-2 |
| 3 | 96 | 100.0 | 19 | 16 | US-09-215-077A-2 |
| 4 | 96 | 100.0 | 206 | 21 | US-09-760-475-2626 |
| 5 | 96 | 100.0 | 347 | 21 | US-09-760-475-2628 |
| 6 | 96 | 100.0 | 383 | 11 | US-08-755-051-3 |
| 7 | 96 | 100.0 | 383 | 11 | US-09-171-562-2 |
| 8 | 96 | 100.0 | 383 | 18 | US-09-459-749B-17 |
| 9 | 96 | 100.0 | 404 | 1 | PCT-US01-14827-14283 |
| | | | | | Sequence 14283, A |

| | | | | | | |
|----|----|------|-----|----|----------------------|-------------------|
| 10 | 85 | 88.5 | 362 | 1 | PCT-US01-14827-14282 | Sequence 14282, A |
| 11 | 69 | 71.9 | 399 | 1 | PCT-US00-23802B-11 | Sequence 11, Appl |
| 12 | 69 | 71.9 | 399 | 1 | PCT-US00-23802B-19 | Sequence 19, Appl |
| 13 | 69 | 71.9 | 399 | 1 | PCT-US00-23802C-11 | Sequence 11, Appl |
| 14 | 69 | 71.9 | 399 | 1 | PCT-US00-23802C-19 | Sequence 19, Appl |
| 15 | 69 | 71.9 | 399 | 20 | US-09-649-747A-11 | Sequence 11, Appl |
| 16 | 69 | 71.9 | 399 | 20 | US-09-649-747A-19 | Sequence 19, Appl |
| 17 | 68 | 70.8 | 188 | 21 | US-09-760-475-2720 | Sequence 17, Appl |
| 18 | 68 | 70.8 | 241 | 1 | PCT-US01-14827-9855 | Sequence 9855, Ap |
| 19 | 68 | 70.8 | 385 | 1 | PCT-US96-13003-12 | Sequence 12, Appl |
| 20 | 68 | 70.8 | 385 | 11 | US-08-755-051-1 | Sequence 1, Appl |
| 21 | 68 | 70.8 | 385 | 15 | US-09-108-346-2 | Sequence 2, Appl |
| 22 | 68 | 70.8 | 416 | 15 | US-09-108-346-4 | Sequence 4, Appl |
| 23 | 68 | 70.8 | 421 | 1 | PCT-US96-13003-2 | Sequence 2, Appl |
| 24 | 68 | 70.8 | 421 | 1 | PCT-US96-13003-4 | Sequence 4, Appl |
| 25 | 68 | 70.8 | 422 | 1 | PCT-US01-14827-9852 | Sequence 9852, Ap |
| 26 | 68 | 70.8 | 438 | 24 | US-60-242-679-1415 | Sequence 1415, Ap |
| 27 | 68 | 70.8 | 451 | 21 | US-09-760-475-2721 | Sequence 2721, Ap |
| 28 | 65 | 67.7 | 531 | 11 | US-08-763-355-12 | Sequence 12, Appl |
| 29 | 64 | 66.7 | 454 | 16 | US-09-252-691-9411 | Sequence 9411, Ap |
| 30 | 64 | 66.7 | 454 | 16 | US-09-252-691C-9411 | Sequence 9411, Ap |
| 31 | 64 | 66.7 | 476 | 24 | US-60-242-679-1419 | Sequence 1419, Ap |
| 32 | 64 | 66.7 | 623 | 11 | US-08-763-355-7 | Sequence 7, Appl |
| 33 | 64 | 66.7 | 678 | 11 | US-08-763-355-2 | Sequence 2, Appl |
| 34 | 64 | 66.7 | 678 | 22 | US-09-876-877-33 | Sequence 33, Appl |
| 35 | 64 | 66.7 | 721 | 11 | US-08-763-355-10 | Sequence 10, Appl |
| 36 | 63 | 65.6 | 158 | 24 | US-60-207-216-836 | Sequence 836, App |
| 37 | 63 | 65.6 | 373 | 10 | US-08-663-618A-14 | Sequence 14, Appl |
| 38 | 63 | 65.6 | 373 | 10 | US-08-663-618A-15 | Sequence 15, Appl |
| 39 | 63 | 65.6 | 373 | 12 | US-08-877-599-14 | Sequence 14, Appl |
| 40 | 63 | 65.6 | 373 | 12 | US-08-877-599-15 | Sequence 15, Appl |
| 41 | 63 | 65.6 | 373 | 16 | US-09-267-574-14 | Sequence 14, Appl |
| 42 | 63 | 65.6 | 373 | 16 | US-09-267-574-15 | Sequence 15, Appl |
| 43 | 63 | 65.6 | 464 | 1 | PCT-US97-05072-2 | Sequence 2, Appl |
| 44 | 63 | 65.6 | 464 | 12 | US-08-819-791-2 | Sequence 2, Appl |
| 45 | 63 | 65.6 | 466 | 1 | PCT-US97-05072-4 | Sequence 4, Appl |

ALIGNMENTS

RESULT 1
US-08-089-989-2
; Sequence 2, Application US/08089989
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.
; APPLICANT: JOHANSEN, JULIA S.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLER HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/089,989
; FILING DATE: 09-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: PD2759
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110

```
;
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 19 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: YKL-40 INTERNAL PEPTIDE A
; FEATURE:
;   NAME/KEY: Peptide
;   LOCATION: 1..19
;
US-08-089-989-2

Query Match      100.0%; Score 96; DB 4; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTLKNRNPMLKTLSSVGG 19
   |||||
Db 1 LNTLKNRNPMLKTLSSVGG 19

RESULT 2
US-09-164-862B-2
; Sequence 2, Application US/09164862B
; GENERAL INFORMATION:
; APPLICANT: Price, Paul
; TITLE OF INVENTION: YKL-40 AS A MARKER AND PROGNOSTIC INDICATOR FOR CANCERS
; FILE REFERENCE: 407T-8955-00US
; CURRENT APPLICATION NUMBER: US/09/164,862B
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-164-862B-2

Query Match      100.0%; Score 96; DB 15; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTLKNRNPMLKTLSSVGG 19
   |||||
Db 1 LNTLKNRNPMLKTLSSVGG 19

RESULT 3
US-09-215-077A-2
; Sequence 2, Application US/09215077A
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; FILE REFERENCE: 407T-89541IUS
; CURRENT APPLICATION NUMBER: US/09/215,077A
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/581,527
; PRIOR FILING DATE: 1996-04-17
; PRIOR APPLICATION NUMBER: 08/089,989
; PRIOR FILING DATE: 1993-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence

;
; INFORMATION: Description of Artificial Sequence: An internal
; OTHER INFORMATION: amino acid sequence for the YKL-40 protein (YKL-40
; OTHER INFORMATION: peptide A).
;
US-09-215-077A-2

Query Match      100.0%; Score 96; DB 16; Length 19;
Best Local Similarity 100.0%; Pred. No. 7.6e-09;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTLKNRNPMLKTLSSVGG 19
   |||||
Db 1 LNTLKNRNPMLKTLSSVGG 19

RESULT 4
US-09-760-475-2626
; Sequence 2626, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2626
; LENGTH: 206
; TYPE: PRT
; ORGANISM: Homo sapiens
;
US-09-760-475-2626

Query Match      100.0%; Score 96; DB 21; Length 206;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 LNTLKNRNPMLKTLSSVGG 19
   |||||
Db 101 LNTLKNRNPMLKTLSSVGG 119

RESULT 5
US-09-760-475-2628
; Sequence 2628, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2628
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (109)
; OTHER INFORMATION: xaa equals any of the naturally occurring L-amino acids
;
US-09-760-475-2628

Query Match      100.0%; Score 96; DB 21; Length 347;
Best Local Similarity 100.0%; Pred. No. 2.7e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

[illegible]

Tue Dec 11 08:46:45 2001

us-09-164-862b-2.rapm

```
; NAME/KEY: DOMAIN
; LOCATION: (43)..(379)
; OTHER INFORMATION: Glycosyl hydrolases family 18 domain identified by Pfam.
; OTHER INFORMATION: accession name Glyco_hydro_18, E-value=3.1e-187, Pfam score of
; OTHER INFORMATION: 630.9
PCT-US01-14827-14283

Query Match      100.0%; Score 96; DB 1; Length 404;
Best Local Similarity 100.0%; Pred. No. 3.2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSSVGG 19
   ||||| ||||| ||||| |||||
Db 101 LNTLKNRPNLKTLSSVGG 119

RESULT 10
PCT-US01-14827-14282
; Sequence 14282, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 14282
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(336)
; OTHER INFORMATION: Glycosyl hydrolases family 18 domain identified by Pfam,
; OTHER INFORMATION: accession name Glyco_hydro_18, E-value=9.3e-184, Pfam score of
; OTHER INFORMATION: 619.2
PCT-US01-14827-14282

Query Match      88.5%; Score 85; DB 1; Length 362;
Best Local Similarity 89.5%; Pred. No. 2.1e-05;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSSVGG 19
   ||||| ||||| ||||| |||||
Db 59 LNTLNTNPNLKTLSSVGG 77

RESULT 11
PCT-US00-23802B-11
; Sequence 11, Application PC/TUS0023802B
; GENERAL INFORMATION:
; APPLICANT: Okubara, Patricia A.
; APPLICANT: Blechl, Ann E.
; APPLICANT: Hohn, Thomas M.
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Cell Wall-Degrading
; TITLE OF INVENTION: Enzymes and Use to Engineer Resistance to Fusarium and
; TITLE OF INVENTION: Other Pathogens
; FILE REFERENCE: 79,99R
; CURRENT APPLICATION NUMBER: PCT/US00/23802B
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,582
; PRIOR FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: 60/224,946
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
```

```
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Fusarium venenatum
PCT-US00-23802B-11

Query Match      71.9%; Score 69; DB 1; Length 399;
Best Local Similarity 73.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSSVGG 19
   ||||| ||||| ||||| |||||
Db 83 LNLKRRNRNLKVLSSIG 101

RESULT 12
PCT-US00-23802B-19
; Sequence 19, Application PC/TUS0023802B
; GENERAL INFORMATION:
; APPLICANT: Okubara, Patricia A.
; APPLICANT: Blechl, Ann E.
; APPLICANT: Hohn, Thomas M.
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Cell Wall-Degrading
; TITLE OF INVENTION: Enzymes and Use to Engineer Resistance to Fusarium and
; TITLE OF INVENTION: Other Pathogens
; FILE REFERENCE: 79,99R
; CURRENT APPLICATION NUMBER: PCT/US00/23802B
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,582
; PRIOR FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: 60/224,946
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Fusarium venenatum
PCT-US00-23802B-19

Query Match      71.9%; Score 69; DB 1; Length 399;
Best Local Similarity 73.7%; Pred. No. 0.012;
Matches 14; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSSVGG 19
   ||||| ||||| ||||| |||||
Db 83 LNLKRRNRNLKVLSSIG 101

RESULT 13
PCT-US00-23802C-11
; Sequence 11, Application PC/TUS0023802C
; GENERAL INFORMATION:
; APPLICANT: Okubara, Patricia A.
; APPLICANT: Blechl, Ann E.
; APPLICANT: Hohn, Thomas M.
; APPLICANT: Berka, Randy M.
; TITLE OF INVENTION: Nucleic Acid Sequences Encoding Cell Wall-Degrading
; TITLE OF INVENTION: Enzymes and Use to Engineer Resistance to Fusarium and
; TITLE OF INVENTION: Other Pathogens
; FILE REFERENCE: 79,99R
; CURRENT APPLICATION NUMBER: PCT/US00/23802C
; CURRENT FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/151,582
; PRIOR FILING DATE: 1999-08-30
; PRIOR APPLICATION NUMBER: 60/224,946
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 399
```


us-09-164-862b-2.rapm

Tue Dec 11 08:46:45 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:35:57 : Search time 169.18 seconds
(without alignments)
1.136 Million cell updates/sec

Title: US-09-164-862B-2

Perfect score: 96

Sequence: 1 LNTLKNRPNLKTLISVGG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 56692 seqs, 10114640 residues

Total number of hits satisfying chosen parameters: 56692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending_Patents_AA_New.*

- 1: /cgn2.6/ptodata/2/paa/PCT_NEW_COMB.pep.*
- 2: /cgn2.6/ptodata/2/paa/US05_NEW_COMB.pep.*
- 3: /cgn2.6/ptodata/2/paa/US07_NEW_COMB.pep.*
- 4: /cgn2.6/ptodata/2/paa/US08_NEW_COMB.pep.*
- 5: /cgn2.6/ptodata/2/paa/US09_NEW_COMB.pep.*
- 6: /cgn2.6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 68 | 70.8 | 423 | 4 | US-08-850-348-2 |
| 2 | 41 | 42.7 | 157 | 5 | US-09-620-394B-8129 |
| 3 | 41 | 42.7 | 180 | 5 | US-09-620-394B-8128 |
| 4 | 41 | 42.7 | 183 | 5 | US-09-620-394B-8127 |
| 5 | 40 | 41.7 | 1765 | 5 | US-09-646-224A-2 |
| 6 | 39 | 40.6 | 580 | 5 | US-09-897-516-6115 |
| 7 | 39 | 40.6 | 627 | 5 | US-09-815-242-13467 |
| 8 | 39 | 40.6 | 630 | 5 | US-09-801-368-220 |
| 9 | 39 | 40.6 | 630 | 5 | US-09-487-558-220 |
| 10 | 39 | 40.6 | 745 | 5 | US-09-815-242-11498 |
| 11 | 38 | 39.6 | 231 | 5 | US-09-731-126-8 |
| 12 | 38 | 39.6 | 275 | 5 | US-09-756-551A-4 |
| 13 | 38 | 39.6 | 500 | 5 | US-09-968-355-26 |
| 14 | 38 | 39.6 | 515 | 5 | US-09-968-355-23 |
| 15 | 38 | 39.6 | 531 | 5 | US-09-968-355-17 |
| 16 | 38 | 39.6 | 583 | 5 | US-09-968-355-17 |
| 17 | 38 | 39.6 | 826 | 5 | US-09-897-516-6831 |
| 18 | 38 | 39.6 | 1057 | 5 | US-09-815-242-5798 |
| 19 | 38 | 39.6 | 1107 | 5 | US-09-815-242-12815 |
| 20 | 38 | 39.6 | 1198 | 5 | US-09-815-242-12446 |
| 21 | 38 | 39.6 | 2384 | 5 | US-09-897-516-4692 |
| 22 | 37.5 | 39.1 | 500 | 5 | US-09-897-516-6875 |
| 23 | 37 | 38.5 | 51 | 5 | US-09-828-792-995 |
| 24 | 37 | 38.5 | 247 | 5 | US-09-620-394B-5275 |
| 25 | 37 | 38.5 | 272 | 5 | US-09-620-394B-5274 |
| 26 | 37 | 38.5 | 316 | 5 | US-09-886-055-39 |
| 27 | 37 | 38.5 | 316 | 5 | US-09-804-291-39 |

| | | | | | | |
|----|----|------|------|---|---------------------|-------------------|
| 28 | 37 | 38.5 | 368 | 5 | US-09-620-394B-5273 | Sequence 5273, Ap |
| 29 | 37 | 38.5 | 488 | 5 | US-09-815-242-13245 | Sequence 13245, A |
| 30 | 37 | 38.5 | 623 | 5 | US-09-897-516-5247 | Sequence 5247, Ap |
| 31 | 37 | 38.5 | 2329 | 5 | US-09-897-516-4693 | Sequence 4693, Ap |
| 32 | 37 | 38.5 | 2890 | 5 | US-09-815-242-11410 | Sequence 11410, A |
| 33 | 37 | 38.5 | 2890 | 5 | US-09-815-242-11571 | Sequence 11571, A |
| 34 | 36 | 37.5 | 267 | 5 | US-09-815-242-5856 | Sequence 5856, Ap |
| 35 | 36 | 37.5 | 269 | 5 | US-09-815-242-12871 | Sequence 12871, A |
| 36 | 36 | 37.5 | 353 | 5 | US-09-815-242-11273 | Sequence 11273, A |
| 37 | 36 | 37.5 | 355 | 5 | US-09-620-394B-6834 | Sequence 6834, Ap |
| 38 | 36 | 37.5 | 437 | 5 | US-09-620-394B-6833 | Sequence 6833, Ap |
| 39 | 36 | 37.5 | 443 | 5 | US-09-976-594-527 | Sequence 527, App |
| 40 | 36 | 37.5 | 453 | 5 | US-09-620-394B-6832 | Sequence 6832, Ap |
| 41 | 36 | 37.5 | 496 | 5 | US-09-815-242-10714 | Sequence 10714, A |
| 42 | 36 | 37.5 | 750 | 5 | US-09-815-242-13405 | Sequence 13405, A |
| 43 | 36 | 37.5 | 1203 | 5 | US-09-989-687-5 | Sequence 5, Appl |
| 44 | 35 | 36.5 | 38 | 5 | US-09-897-516-5428 | Sequence 5428, Ap |
| 45 | 35 | 36.5 | 97 | 5 | US-09-703-101-21 | Sequence 21, Appl |

ALIGNMENTS

RESULT 1
US-08-850-348-2
; Sequence 2, Application US/08850348
; GENERAL INFORMATION:
; APPLICANT: Kirkpatrick, Robert
; Rosenberg, Martin
; TITLE OF INVENTION: Human Cartilage Glycoprotein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SmithKline Beecham Corporation
; STREET: 709 Swedeland Road
; CITY: King of Prussia
; STATE: PA
; COUNTRY: USA
; ZIP: 19406-0939
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION NUMBER: US/08/850,348
; FILING DATE: 02-May-1997
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/016,532
; FILING DATE: 03-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Han, William T
; REGISTRATION NUMBER: 5219
; REFERENCE/DOCKET NUMBER: P50390
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-270-5219
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 423 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-08-850-348-2

Query Match 70.8%; Score 68; DB 4; Length 423;
Best Local Similarity 63.2%; Pred. No. 0.00063;
Matches 12; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Qy 1 LNTLKNRPNLKTLSSVG 19
: : : : : : : : : :
Db 118 INSLTKNPKLKILLSIG 136

RESULT 2
US-09-620-394B-8129
. sequence 8129. Application US/09620394B

; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai

APPLICANT: BROVER, Vyacheslav
TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
DATE OF INVENTION: Thereby

FILE OF INVENTION: 2750-1067P
CURRENT APPLICATION NUMBER: US/09/620,394B
CURRENT FILING DATE: 2000-07-21

```

; CURRENT FILING DATE: 2000 07 21
; NUMBER OF SEQ ID NOS: 9131

```

; SEQ ID NO 8129

```

; LENGTH: 157

```

```

;
; TYPE: PRT

```

; ORGANISM: Arabidopsis thaliana

```

; FEATURE:
; NAME/VEY. misc feature

```

```

; NAME/KEY: misc_feature
; LOCATION: 1 157

```

LOCATION: 1..157
OTHER INFORMATION: Xaa is any amino acid

```

; OTHER INFORMATION:
; NAME/KEY: misc_feature

```

LOCATION: 1..157

OTHER INFORMATION: Ceres Seq. ID 1411643

US-09-620-394B-8129

Query Match 42.7%; Score 41; DB 5; Length 157;

Best Local Similarity 53.3%; Pred. No. 5.2;

| | | | | | | | | | |
|----------------|-----------|---------------------|-----------|-------------------|-----------|---------------|-----------|-------------|-----------|
| Matches | 8; | Conservative | 4; | Mismatches | 3; | Indels | 0; | Gaps | 0; |
|----------------|-----------|---------------------|-----------|-------------------|-----------|---------------|-----------|-------------|-----------|

0v 4 LKRNPNLKTLLSVG 18

[illegible]

RESULT 3

US-09-620-394B-8128

; Sequence 8128, Application US/09620394B

GENERAL INFORMATION:

APPLICANT: ALEXANDROV, Nikolai

; APPLICANT: BROVER, Vyacheslav
 ; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
 ; TITLE OF INVENTION: Thereby

FILE REFERENCE: 2750-1067P

; CURRENT APPLICATION NUMBER: US/09/620,394B

; CURRENT FILING DATE: 2000-07-21
NUMBER OF SEC ID NOS: 8131

```

; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 8128

```

```
; SEQ ID NO 8128
:      LENGTH: 180
```

```

; LENGTH: 100
; TYPE: PRT
;

```

ORGANISM: *Arabidopsis thaliana*

FEATURE:

```
; NAME/KEY: misc_fe
```

; LOCATION: 1.180

```
; OTHER INFORMATION: Xaa
: NAME/KEY: misc feature
```

```

; NAME/KEY: MISC_F
; LOCATION: 1-180

```

LOCATION: 1.1.180
OTHER INFORMATION: Ceres Seq. ID 1411642

US-09-620-394B-8128

Query Match 42.7% Score 41; DB 5; Length 180;

Query Match

| Best Local Similarity Score, α | Mismatches | Conservative | Mismatches | Indels | Gaps |
|---------------------------------------|------------|--------------|------------|--------|------|
| 0.0 | 8 | 8 | 4 | 3 | 0 |
| 0.1 | 8 | 8 | 4 | 3 | 0 |
| 0.2 | 8 | 8 | 4 | 3 | 0 |
| 0.3 | 8 | 8 | 4 | 3 | 0 |
| 0.4 | 8 | 8 | 4 | 3 | 0 |
| 0.5 | 8 | 8 | 4 | 3 | 0 |
| 0.6 | 8 | 8 | 4 | 3 | 0 |
| 0.7 | 8 | 8 | 4 | 3 | 0 |
| 0.8 | 8 | 8 | 4 | 3 | 0 |
| 0.9 | 8 | 8 | 4 | 3 | 0 |
| 1.0 | 8 | 8 | 4 | 3 | 0 |

Q17 A I.K.NRNPNT.K.TT.I.SVG 18

QY 4 LKRNPNLNLLESG TO
| | : | | : | | : |

Db 134 LLNKSPNLQTLVVG 148

; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesing, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 2001-06-29
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 6115
; LENGTH: 580
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516-6115

Query Match 40.6%; Score 39; DB 5; Length 580;
Best Local Similarity 43.8%; Pred. No. 53;
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 2 NTLKRNPNLKTLLSV 17

Db 555 NEIQNAKPNLTQLQI 570

RESULT 7

US-09-815-242-13467
; Sequence 13467, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA-011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13467
; LENGTH: 627
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13467

Query Match 40.6%; Score 39; DB 5; Length 627;
Best Local Similarity 63.6%; Pred. No. 59;

Matches 7; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 4 LKNRPNLKTLL 14
Db 347 VRKNPNLKTLL 357

RESULT 8

US-09-801-368-220
; Sequence 220, Application US/09801368
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Norman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Soïe
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 220
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-220

Query Match 40.6%; Score 39; DB 5; Length 630;
Best Local Similarity 58.3%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 TLKRNPNLKTLL 14

Db 323 TINNSNPNEKLL 334

RESULT 9

US-09-487-558-220
; Sequence 220, Application US/09487558
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: Norman, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Soïe
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/487,558
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US/09/801,368

Tue Dec 11 08:46:46 2001

us-09-164-862b-2.rapn

```
; PRIOR FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: US 60/160,587
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 220
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Saccharomyces cerevisiae
US-09-487-558-220

Query Match          40.6%; Score 39; DB 5; Length 630;
Best Local Similarity 58.3%; Pred. No. 59;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 3 TLKRNPNLTKL 14
   |: | | | |
Db 323 TINNSPNFKLL 334

RESULT 10
US-09-815-242-11498
; Sequence 11498, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA 011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11498
; LENGTH: 745
; TYPE: PRT
; ORGANISM: Helicobacter pylori
US-09-815-242-11498

Query Match          40.6%; Score 39; DB 5; Length 745;
Best Local Similarity 40.0%; Pred. No. 72;
Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

QY 4 LKNRNPNLTKLLSVG 18
   |: | | | |
Db 149 LWHQPNMSSLIAIG 163

RESULT 11
US-09-731-126-8
; Sequence 8, Application US/09731126
; GENERAL INFORMATION:
; APPLICANT: Abbott Laboratories
; APPLICANT: Lou, Sheng C.
; APPLICANT: Hunt, Jeffrey C.
; APPLICANT: Konrath, John G.
; APPLICANT: Qiu, Xiaoxing
; APPLICANT: Scheffel, James W.
; APPLICANT: Tyner, Joan D.
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES TO HUMAN
; FILE REFERENCE: 6755.US.01
; CURRENT APPLICATION NUMBER: US/09/731,126
; CURRENT FILING DATE: 2000-12-06
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Human Immunodeficiency Virus
US-09-731-126-8

Query Match          39.6%; Score 38; DB 5; Length 231;
Best Local Similarity 43.8%; Pred. No. 26;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 LKNRNPNLTKLLSVG 19
   |: | | | |
Db 191 VQVNPDCITLKALG 206

RESULT 12
US-09-756-551A-4
; Sequence 4, Application US/09756551A
; GENERAL INFORMATION:
; APPLICANT: C. Morrow et al.
; TITLE OF INVENTION: ENCAPSIDATED RECOMBINANT VIRAL
; TITLE OF INVENTION: NUCLEIC ACID AND METHODS OF MAKING AND
; TITLE OF INVENTION: USING SAME
; NUMBER OF SEQUENCES: 23
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 STATE STREET
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/756,551A
; FILING DATE: 08-JAN-2001
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/376,184
; FILING DATE: 17-AUG-1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/987,867
; FILING DATE: 09-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/389,459
; FILING DATE: 15-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/087,009
; FILING DATE: 01-JUL-1993
; ATTORNEY/AGENT INFORMATION:
```

; NAME: Lauro, Peter C.
; REGISTRATION NUMBER: 32,360
; REFERENCE/DOCKET NUMBER: UAI-004CPDV2CN
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 275 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-756-551A-4

Query Match 39.6%; Score 38; DB 5; Length 275;
Best Local Similarity 43.8%; Pred. No. 32;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 LKNRNPNTLLSVGG 19
::| ||: ||: |
Db 194 VQNPDPCKTILKALG 209

RESULT 13

US-09-968-355-26
; Sequence 26, Application US/09968355
; GENERAL INFORMATION:
; APPLICANT: Sakalian, Michael
; TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays
; FILE REFERENCE: UAB-100XC1
; CURRENT APPLICATION NUMBER: US/09/968,355
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,273
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 500
; TYPE: PRT
; ORGANISM: Human Immunodeficiency virus
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(500)
; OTHER INFORMATION: HIV (FS-) GAG amino acid sequence
US-09-968-355-26

Query Match 39.6%; Score 38; DB 5; Length 500;
Best Local Similarity 43.8%; Pred. No. 65;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 LKNRNPNTLLSVGG 19
::| ||: ||: |
Db 323 VQNPDPCKTILKALG 338

RESULT 14

US-09-968-355-20
; Sequence 20, Application US/09968355
; GENERAL INFORMATION:
; APPLICANT: Sakalian, Michael
; TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays
; FILE REFERENCE: UAB-100XC1
; CURRENT APPLICATION NUMBER: US/09/968,355
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,273
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20

; LENGTH: 515
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(515)
; OTHER INFORMATION: Ch3a amino acid sequence
US-09-968-355-20

Query Match 39.6%; Score 38; DB 5; Length 515;
Best Local Similarity 43.8%; Pred. No. 68;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 LKNRNPNTLLSVGG 19
::| ||: ||: |
Db 323 VQNPDPCKTILKALG 338

RESULT 15

US-09-968-355-23
; Sequence 23, Application US/09968355
; GENERAL INFORMATION:
; APPLICANT: Sakalian, Michael
; TITLE OF INVENTION: Chimeric Retroviral Gag Genes and Screening Assays
; FILE REFERENCE: UAB-100XC1
; CURRENT APPLICATION NUMBER: US/09/968,355
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/236,273
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 23
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(531)
; OTHER INFORMATION: Ch3 amino acid sequence
US-09-968-355-23

Query Match 39.6%; Score 38; DB 5; Length 531;
Best Local Similarity 43.8%; Pred. No. 70;
Matches 7; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 4 LKNRNPNTLLSVGG 19
::| ||: ||: |
Db 323 VQNPDPCKTILKALG 338

Search completed: December 7, 2001, 00:35:57
Job time: 871 sec

us-09-164-862b-2.rapn

Tue Dec 11 08:46:46 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:26:03 ; Search time 71.99 Seconds
(without alignments)
20.104 Million cell updates/sec

Title: US-09-164-862B-2

Perfect score: 96

Sequence: 1 LNTLKNRPNLKTLSSVGG 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR_68:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|------------------------|
| 1 | 96 | 100.0 | 383 | 2 S51327 | heparin-binding |
| 2 | 96 | 100.0 | 383 | 2 A49562 | cartilage glycoprotein |
| 3 | 76 | 79.2 | 405 | 2 S61551 | breast-regressing |
| 4 | 74 | 77.1 | 699 | 2 A38368 | chitinase (EC 3.2. |
| 5 | 67 | 69.8 | 599 | 2 D83764 | chitinase BH0916 [|
| 6 | 65 | 67.7 | 539 | 2 I46470 | estrogen dependent |
| 7 | 64 | 66.7 | 332 | 2 T04754 | hypothetical prote |
| 8 | 64 | 66.7 | 366 | 2 T04763 | chitinase homolog |
| 9 | 64 | 66.7 | 379 | 2 T04762 | chitinase homolog |
| 10 | 64 | 66.7 | 398 | 2 T04761 | chitinase homolog |
| 11 | 64 | 66.7 | 537 | 2 S57197 | oviduct-specific g |
| 12 | 64 | 66.7 | 654 | 2 I38605 | chitinase (EC 3.2. |
| 13 | 60 | 62.5 | 427 | 2 JC4565 | oviductal glycopro |
| 14 | 57 | 59.4 | 1054 | 2 T30933 | chitinase (EC 3.2. |
| 15 | 56 | 58.3 | 421 | 2 T04753 | hypothetical prote |
| 16 | 55 | 57.3 | 371 | 2 T04756 | chitinase homolog |
| 17 | 55 | 57.3 | 452 | 2 JC4038 | 47K glycoprotein p |
| 18 | 54 | 56.2 | 270 | 1 S62837 | protein soj homolo |
| 19 | 53 | 55.2 | 365 | 2 T04757 | chitinase homolog |
| 20 | 53 | 55.2 | 423 | 2 TQ1975 | chitinase (EC 3.2. |
| 21 | 53 | 55.2 | 1088 | 2 D82246 | probable chitinase |
| 22 | 52.5 | 54.7 | 399 | 2 S27879 | secretory protein |
| 23 | 52 | 54.2 | 144 | 2 S71764 | narbonin (clone pN |
| 24 | 52 | 54.2 | 378 | 2 S51591 | chitinase (EC 3.2. |
| 25 | 52 | 54.2 | 424 | 2 S47133 | chitinase (EC 3.2. |
| 26 | 52 | 54.2 | 275 | 2 T44445 | chitinase (EC 3.2. |
| 27 | 51 | 53.1 | 226 | 2 S56696 | narbonin (clone pN |
| 28 | 51 | 53.1 | 291 | 2 S56697 | narbonin (clone pN |
| 29 | 51 | 53.1 | 291 | 2 S56698 | narbonin (clone pN |

ALIGNMENTS

RESULT 1

S51327

heparin-binding glycoprotein 38K - pig

C:Species: Sus scrofa domestica (domestic pig)

C>Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999

C:Accession: S51327

R:Shackelton, L.M.; Mann, D.M.; Mills, A.J.T.

submitted to the EMBL Data Library, January 1995

A:Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in diff

A:Reference number: S51327

A:Accession: S51327

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-383 <SHA>

A:Cross-references: EMBL:47803; NID:g634097; PIDN:CAA87764.1; PID:g634098

C:Superfamily: Streptomyces chitinase chi40

Query Match 100.0%; Score 96; DB 2; Length 383;

Best Local Similarity 100.0%; Pred. NO. 1.2e-07;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSSVGG 19

Db 80 LNTLKNRPNLKTLSSVGG 98

RESULT 2

A49562

cartilage glycoprotein gp39 precursor - human

N:Alternate names: 39K synovial protein

C:Species: Homo sapiens (man)

C>Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999

C:Accession: A49562; S10677; A33162

R:Hakala, B.E.; White, C.; Recklies, A.D.

J. Biol. Chem. 268, 25803-25810, 1993

A:Title: Human cartilage gp-39, a major secretory product of articular chondrocytes a

A:Reference number: A49562; MUID:94064658

A:Accession: A49562

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-383 <HAK>

A:Cross-references: GB:M80927; NID:g348911; PIDN:AAA16074.1; PID:g348912

R:Nyrkos, P.; Golds, E.E.

Biochem. J. 269, 265-268, 1990

A:Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary pr

A:Reference number: S10677; MUID:90328983

A:Accession: S10677

A:Molecule type: protein

A:Residues: 22-40,'x',42-45 <NY2>

C:Superfamily: Streptomyces chitinase chi40

C:Keywords: cartilage; extracellular protein; glycoprotein

| | | | | | | |
|----|----|------|-----|---|--------|--------------------|
| 30 | 51 | 53.1 | 499 | 2 | S52422 | chitinase (EC 3.2. |
| 31 | 51 | 53.1 | 499 | 2 | S04856 | chitinase (EC 3.2. |
| 32 | 51 | 53.1 | 799 | 2 | PC4106 | chitinase (EC 3.2. |
| 33 | 51 | 53.1 | 820 | 2 | A40633 | chitinase (EC 3.2. |
| 34 | 51 | 53.1 | 869 | 2 | T44440 | chitinase (EC 3.2. |
| 35 | 50 | 52.1 | 213 | 2 | H64484 | hypothetical prote |
| 36 | 50 | 52.1 | 504 | 2 | A38221 | chitinase (EC 3.2. |
| 37 | 48 | 50.0 | 117 | 2 | T37249 | chitinase (EC 3.2. |
| 38 | 48 | 50.0 | 124 | 2 | S57715 | chitinase (EC 3.2. |
| 39 | 48 | 50.0 | 285 | 2 | S49879 | hypothetical narbo |
| 40 | 48 | 50.0 | 285 | 2 | S49898 | hypothetical narbo |
| 41 | 48 | 50.0 | 483 | 2 | A53918 | chitinase (EC 3.2. |
| 42 | 48 | 50.0 | 617 | 2 | T15408 | hypothetical prote |
| 43 | 48 | 50.0 | 831 | 2 | T00323 | chitinase (EC 3.2. |
| 44 | 47 | 49.0 | 285 | 2 | T12156 | nodulin, isoform N |
| 45 | 47 | 49.0 | 285 | 2 | T12157 | nodulin - fava bea |

Tue Dec 11 08:46:46 2001

us-09-164-862b-2.rpr

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-383/Product: cartilage glycoprotein gp39 #status predicted <MAT>

Query Match 100.0%; Score 96; DB 2; Length 383;
Best Local Similarity 100.0%; Pred. No. 1.2e-07;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTLKNRNPMLKTLISVGG 19
|||||
Db 80 LNTLKNRNPMLKTLISVGG 98

RESULT 3

S61551 breast-regressing protein brp39 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000

C:Accession: S61551; S61550; I48271

R:Morrison, B.W.; Leder, P.

Oncogene 9, 3417-3426, 1994

A:Title: neu and ras initiate murine mammary tumors that share genetic markers generally

A:Reference number: I48271; MUID:95060797

A:Accession: S61551

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-405 <MOR1>

A:Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066

R:Morrison, B.W.

submitted to the EMBL Data Library, November 1995

A:Accession: S61550

A:Molecule type: mRNA

A:Residues: 1-245, 'I', 247-330, 'H', 332-350, 'MVWALDLDFOGTCQPKFFPLTNAIKDALA' <MOR2>

A:Cross-references: EMBL:X93035; NID:g1085065; PIDN:CAA63603.1; PID:g1085066

A:Note: the differences at the carboxyl end are due to a frameshift error

C:Genetics:

A:Gene: brp39

C:Superfamily: Streptomyces chitinase chl40

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-405/Product: breast-regressing protein brp39 #status predicted <MAT>

Query Match 79.2%; Score 76; DB 2; Length 405;
Best Local Similarity 84.2%; Pred. No. 0.0002;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNTLKNRNPMLKTLISVGG 19
|||||
Db 81 LNKLTNTNLTLLSVGG 99

RESULT 4

A38368 chitinase (EC 3.2.1.14) precursor - Bacillus circulans

C:Species: Bacillus circulans

C:Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 15-Oct-1999

R:Watanabe, T.; Suzuki, K.; Oyanagi, W.; Ohnishi, K.; Tanaka, H.

J. Biol. Chem. 265, 15659-15665, 1990

A:Title: Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed its evolution

A:Reference number: A38368; MUID:90368776

A:Accession: A38368

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-699 <MAT>

A:Cross-references: GB:M57601; GB:J05599; NID:g1066341; PIDN:AAA81528.1; PID:g142688

C:Superfamily: fibronectin type III repeat homology

C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 77.1%; Score 74; DB 2; Length 699;
Best Local Similarity 73.7%; Pred. No. 0.00079;

A:Accession: T04754

A:Reference number: 215383

A:Accession: T04754

A:Reference number: 215383

A:Accession: T04754

A:Reference number: 215383

A:Accession: T04754

Tue Dec 11 08:46:46 2001

us-09-164-862b-2.rpr

```
A:Reference number: I38605; MUID:95119256
A:Accession: I38605
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-654 <RES>
A:Cross-references: EMBL:U09550; NID:g529147; PID:g529148

Query Match      66.7%; Score 64; DB 2; Length 654;
Best Local Similarity 72.2%; Pred. No. 0.029;
Matches 13; Conservative 1; Mismatches 0; Gaps 0;

QY  2 NTLKRNPNLKTLLSVGG 19
    |||||
Db   82 NTLKRNPNLKTLLSVGG 99
    |||||

RESULT 13
JC4565
Chitinase (EC 3.2.1.14) 1 precursor - Coccidioides immitis
N:Alternate names: complement fixation antigen homolog; CTs1 protein
C:Species: Coccidioides immitis
C:Date: 12-Mar-1996 #sequence_revision 19-Apr-1996 #text_change 13-Nov-1998
C:Accession: JC4565
R:Pishko, E.J.; Kirkland, T.N.; Cole, G.T.
Gene 167, 173-177, 1995
A:Title: Isolation and characterization of two chitinase-encoding genes (ctsl, cts2) from
A:Reference number: JC4565; MUID:96144270
A:Accession: JC4565
A:Molecule type: mRNA
A:Residues: 1-427 <PIS>
A:Cross-references: GB:L41663
A:Experimental source: C735
C:Genetics:
A:Gene: cts1
A:Introns: 47/3; 171/3; 191/3; 215/2; 393/3
C:Superfamily: Streptomyces chitinase chi40
C:Keywords: glycoprotein; glycosidase; hydrolase; polysaccharide degradation
F:1-17/Domain: signal sequence #status predicted <SIG>
F:18-427/Product: chitinase 1 #status predicted <MAT>
F:387/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match      62.5%; Score 60; DB 2; Length 427;
Best Local Similarity 75.0%; Pred. No. 0.08;
Matches 12; Conservative 1; Mismatches 0; Gaps 0;

QY  4 LKNRNPNLKTLLSVGG 19
    |||||
Db  115 LKNRNPNLKTLLSVGG 130
    |||||

RESULT 14
T30933
chitinase (EC 3.2.1.14) A - Pseudoalteromonas sp. (strain S9)
C:Species: Pseudoalteromonas sp.
A:Variety: strain S9
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 02-Sep-2000
C:Accession: T30933
R:Techkarnjanaruk, S.; Goodman, A.E.
Microbiology 145, 925-934, 1999
A:Title: Multiple genes involved in chitin degradation from the marine bacterium Pseudoa
A:Reference number: 220935; MUID:99235578
A:Accession: T30933
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1054 <TEC>
A:Cross-references: EMBL:AF007894; NID:g3928771; PID:g3928775; PIDN:RAC79665.1
C:Genetics:
A:Gene: chiA
C:Keywords: glycosidase; hydrolase
```

```
Query Match      59.4%; Score 57; DB 2; Length 1054;
Best Local Similarity 57.9%; Pred. No. 0.67;
Matches 11; Conservative 4; Mismatches 0; Gaps 0;

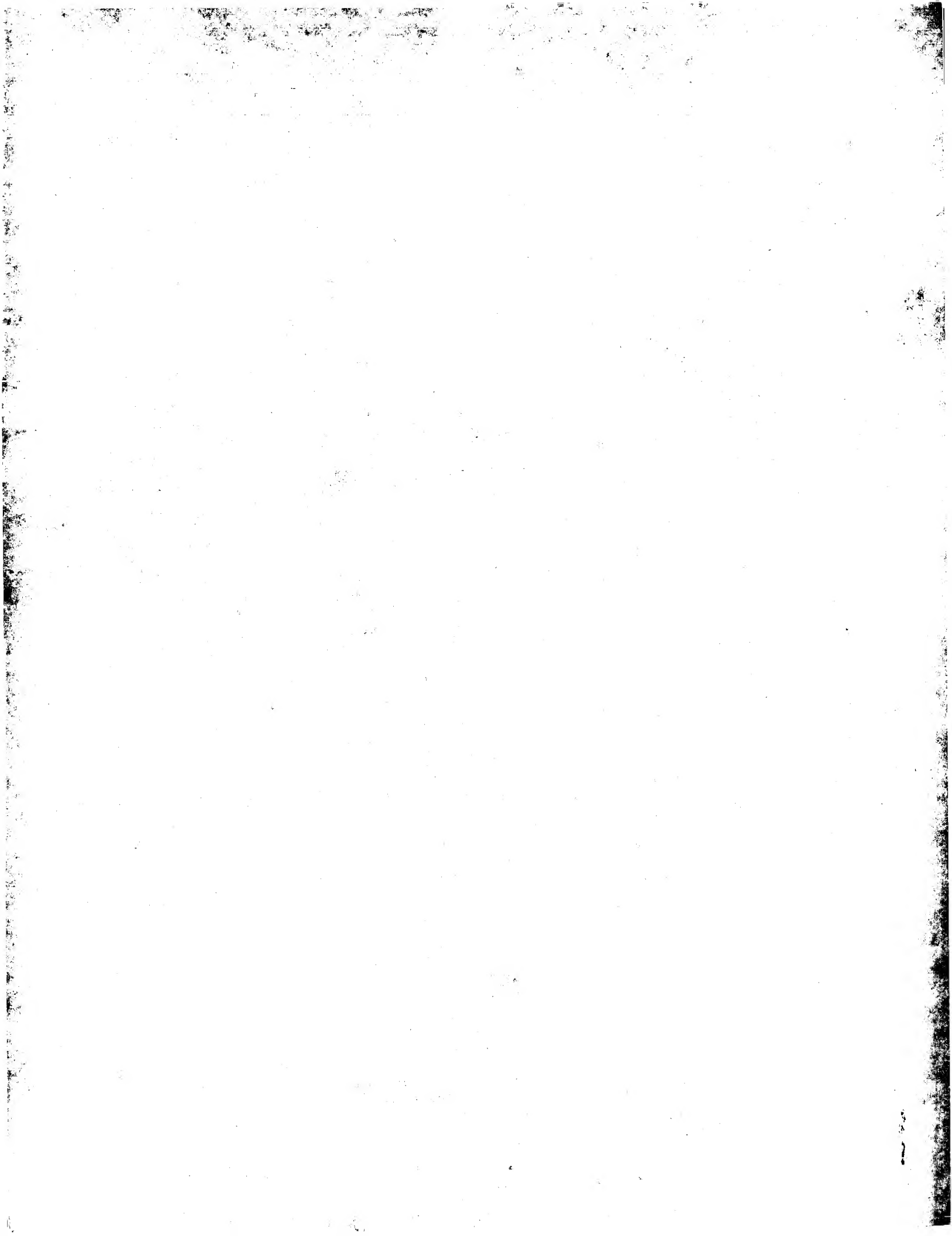
QY  1 LNTLKNRNPNLKTLLSVGG 19
    |||||
Db   403 LNKYKLLHPDVKTLLSVGG 421
    |||||

RESULT 15
T04753
hypothetical protein T16H5.80 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 26-Aug-1999
C:Accession: T04753
R:Bevan, M.; De Haan, M.; Maarse, A.C.; Grivell, L.A.; Bancroft, I.; Mewes, H.W.; May
submitted to the Protein Sequence Database, June 1998
A:Reference number: Z15383
A:Accession: T04753
A:Molecule type: DNA
A:Residues: 1-421 <BEV>
A:Cross-references: EMBL:AL024486
A:Experimental source: cultivar Columbia; BAC clone T16H5
C:Genetics:
A:Map position: 4
A:Introns: 278/1
A:Note: T16H5.80
C:Superfamily: Streptomyces chitinase chi40

Query Match      58.3%; Score 56; DB 2; Length 421;
Best Local Similarity 56.2%; Pred. No. 0.35;
Matches 9; Conservative 6; Mismatches 1; Indels 0; Gaps 0;

QY  4 LKNRNPNLKTLLSVGG 19
    |||||
Db  72 VKKNPHVQTLISIGG 87
    |||||

Search completed: December 7, 2001, 00:26:04
Job time: 2844 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

| Result No. | Score | Query % | | DB | ID | Description |
|------------|-------|---------|--------|----|------------|---------------------|
| | | Match | Length | | | |
| 1 | 96 | 100.0 | 383 | 1 | C3L1_HUMAN | P36222 homo sapien |
| 2 | 76 | 79.2 | 381 | 1 | C3L1_MOUSE | O61362 mus musculus |
| 3 | 74 | 77.1 | 699 | 1 | CH11_BACCI | P20533 bacillus ci |
| 4 | 68 | 70.8 | 390 | 1 | C3L2_HOMAN | Q15782 homo sapien |
| 5 | 65 | 67.7 | 527 | 1 | OGP_PIG | Q28990 sus scrofa |
| 6 | 65 | 67.7 | 539 | 1 | OGP_SHEEP | Q28542 ovis aries |
| 7 | 64 | 66.7 | 537 | 1 | OGP_BOVIN | Q28042 bos taurus |
| 8 | 64 | 66.7 | 623 | 1 | OGP_PAPAN | P36718 papio anubi |
| 9 | 64 | 66.7 | 678 | 1 | OGP_HUMAN | Q12889 homo sapien |
| 10 | 64 | 66.7 | 721 | 1 | OGP_MOUSE | Q62010 mus musculus |
| 11 | 63 | 65.6 | 671 | 1 | OGP_MESAU | Q60557 mesocricetu |
| 12 | 60 | 62.5 | 427 | 1 | CH11_COCIM | P54196 coccidioide |
| 13 | 54 | 56.2 | 270 | 1 | PARA_MYCPN | Q50314 mycoplasma |
| 14 | 53 | 55.2 | 423 | 1 | CH11_APHAL | P32470 apyanoclad |
| 15 | 51 | 53.1 | 499 | 1 | CH1B_SERMA | P11797 serratia ma |
| 16 | 51 | 53.1 | 820 | 1 | CH1A_ALTSO | P32823 alteromonas |
| 17 | 50 | 52.1 | 213 | 1 | YES1_METJA | Q58876 methanococc |
| 18 | 50 | 52.1 | 504 | 1 | CH1T_BRUMA | P29030 bugia mala |
| 19 | 48 | 50.0 | 617 | 1 | CH1T_CAEEL | Q11174 caenorhabdi |
| 20 | 47 | 49.0 | 563 | 1 | CH1A_SERMA | P27254 serratia ma |
| 21 | 47 | 49.0 | 619 | 1 | CH1T_STRLI | P36909 streptomyce |
| 22 | 46 | 47.9 | 269 | 1 | PARA_MYCGE | P47706 mycoplasma |
| 23 | 45 | 46.9 | 218 | 1 | Y010_MYCGE | P47256 mycoplasma |
| 24 | 45 | 46.9 | 309 | 1 | MAT1_HUMAN | P51948 homo sapien |
| 25 | 45 | 46.9 | 309 | 1 | MAT1_MOUSE | P51949 mus musculus |
| 26 | 45 | 46.9 | 309 | 1 | MAT1_XENLA | P51951 xenopus lae |
| 27 | 45 | 46.9 | 355 | 1 | BUK_CLOBE | Q05619 clostridium |
| 28 | 45 | 46.9 | 423 | 1 | CH14_TRIHA | P48827 trichoderma |
| 29 | 44 | 45.8 | 550 | 1 | CH1T_NPVOP | O10363 orgyia pseu |
| 30 | 44 | 45.8 | 551 | 1 | CH1T_NPVAC | P41684 autographa |
| 31 | 43 | 44.8 | 292 | 1 | CH1A_CUCSA | P17541 cucumis sat |
| 32 | 43 | 44.8 | 292 | 1 | PPNK_BUCAI | P57282 buchnera ap |
| 33 | 43 | 44.8 | 324 | 1 | MAT1_MARGL | P51950 marthasteri |

Tue Dec 11 08:46:47 2001

```
DR EMBL; M80927; AAA16074.1; -
DR EMBL; Y08374; CAA69661.1; -
DR EMBL; Y08375; CAA69661.1; JOINED.
DR EMBL; Y08376; CAA69661.1; JOINED.
DR EMBL; Y08377; CAA69661.1; JOINED.
DR EMBL; Y08378; CAA69661.1; JOINED.
DR FIR; S10677; S10677.
DR PIR; A33162; A33162.
DR MIM; 601525; -.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 383
FT CARBOHYD 60 60
FT SEQUENCE 383 AA; 42613 MW; 76ADD8298EEC2D1 CRC64;
SQ
Query Match 100.0%; Score 96; DB 1; Length 383;
Best Local Similarity 100.0%; Pred. No. 5e-08; 0; Indels 0; Gaps 0;
Matches 19; Conservative 0; Mismatches 0;
QY 1 LNTLKNRPNLKTLSSVG 19
DB 80 LNTLKNRPNLKTLSSVG 98
RESULT 2
C3L1_MOUSE
ID C3L1_MOUSE STANDARD; PRT; 381 AA.
AC Q61362; 2001 (Rel. 40, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE CHITINASE-3 LIKE PROTEIN 1 PRECURSOR (CARTILAGE GLYCOPROTEIN-39)
GN (GP-39) (BRP39 PROTEIN).
DN CH3L1 OR BRP39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FVB/N; TISSUE=Breast;
RX MEDLINE=95060797; PubMed=7970700;
RA Morrison B.W., Leder P.;
RT "new and ras initiate murine mammary tumors that share genetic markers
RT generally absent in c-myc and int-2-initiated tumors.";
RL Oncogene 9:3417-3426(1994).
CC -!- FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO
CC RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: EXTRACELLULAR (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X93035; CAA63603.1; -
CC MGD; MGI:1340899; Ch3l1.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 383
FT CARBOHYD 60 60
FT SEQUENCE 381 AA; 43001 MW; EF6581E8184F0450 CRC64;
SQ
Query Match 79.2%; Score 76; DB 1; Length 381;
Best Local Similarity 84.2%; Pred. No. 8.2e-05;
Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 LNTLKNRPNLKTLSSVG 19
DB 81 LNTLKNRPNLKTLSSVG 99
RESULT 3
CHIL_BACCI
ID CHIL_BACCI STANDARD; PRT; 699 AA.
AC P20533;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 01-FEB-1991 (Rel. 31, Last annotation update)
DE CHITINASE A1 PRECURSOR (EC 3.2.1.14).
GN CHIA1.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WL-12;
RX MEDLINE=90368776; PubMed=2203782;
RA Watanabe T., Suzuki K., Oyanagi W., Ohnishi K., Tanaka H.;
RT "Gene cloning of chitinase A1 from Bacillus circulans WL-12 revealed
RT its evolutionary relationship to Serratia chitinase and to the type
RT III homology units of fibronectin.";
RL J. Biol. Chem. 265:15659-15665(1990).
RN [2]
RP MUTAGENESIS.
RC STRAIN=WL-12;
RX MEDLINE=93366760; PubMed=8103047;
RA Watanabe T., Kohori K., Miyashita K., Fujii T., Sakai H.;
RA Uchida M., Tanaka H.;
RT "Identification of glutamic acid 204 and aspartic acid 200 in
RT chitinase A1 of Bacillus circulans WL-12 as essential residues for
RT chitinase activity.";
RL J. Biol. Chem. 268:18567-18572(1993).
CC -!- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF
CC N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN.
CC -!- SIMILARITY: CONTAINS 2 FIBRONECTIN TYPE III-LIKE DOMAINS.
CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M57601; AAA81528.1; -
CC PIR; A38368; A38368.
DR HSSP; P02751; 1TTG.
DR InterPro; IPR003610; Chitin_bind3.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001777; FN.III.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00014; FNTYPEII.
DR SMART; SM00495; ChtBD3; 1.
DR SMART; SM00060; FN3; 2.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
```

DR EMBL; U58515; AAB04534.1; -.
DR EMBL; U58514; AAC04533.1; -.
DR EMBL; U49835; AAC50597.1; ALT_INIT.
DR MIN; 601526; -.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001233; Glyco_hydro_18
DR Pfam; PF00704; Glyco_hydro_18; 1.

Query Match 67.7%; Score 65; DB 1; Length 527;
Best Local Similarity 72.2%; Pred. No. 0.0069;
Matches 13: Conservative 1; Mismatches 4; Indels

```

QY 2 NTLKNRNPNTLLSVGG 19
DB 82 NQKERNRGLKTLISVG 99

RESULT 6
OGP_SHEEP STANDARD; PRT; 539 AA.
AC Q28542; Q28543;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN) (ESTRUS-ASSOCIATED
DE OVIDUCTAL GLYCOPROTEIN) (OEGP).
GN OVGPI OR OGP.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 22-39.
RX TISSUE=oviduct;
RX MEDLINE=95269691; PubMed=7750470;
RA Desouza M.M., Murray M.K.;
RA "An estrogen-dependent secretory protein, which shares identity with
RT chitinases, is expressed in a temporally and regionally specific
RT manner in the sheep oviduct at the time of fertilization and embryo
RT development.";
RT Endocrinology 136:2485-2496(1995).
RN [2]
RP SEQUENCE OF 10-539 FROM N.A.
RC STRAIN=MERINO; TISSUE=oviduct;
RX MEDLINE=96329120; PubMed=8726871;
RA Marshall J.T.A., Nancarrow C.D., Brownlee A.G.;
RA "Cloning and sequencing of a cDNA encoding an ovine
RT oestrus-associated oviductal protein.";
RT Reprod. Fertil. Dev. 8:305-310(1996).
CC -1- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -1- TISSUE SPECIFICITY: OVIDUCT.
CC -1- DEVELOPMENTAL STAGE: LEVELS ARE HIGHEST IN THE FIMBRIA AND AMPULLA
CC AT ESTRUS AND ON DAY 1 OF PREGNANCY, WHEN GAMETE TRANSPORT AND
CC FERTILIZATION OCCURS IN THE E2-DOMINATED FALLOPIAN TUBE. LEVELS
CC DECLINE SIGNIFICANTLY ON DAY 2 AND UNDERGO A FURTHER SIGNIFICANT
CC REDUCTION ON DAY 3 OF PREGNANCY COINCIDENT WITH TRANSPORT OF THE
CC EMBRYO FROM THE OVIDUCT TO THE UTERUS, A REPRODUCTIVE STAGE
CC ASSOCIATED WITH RISING PROGESTERONE LEVELS.
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U16719; AAC48471.1; -.
CC EMBL; U17988; AAB01052.1; -.
CC InterPro: IPR001579; Chitinase_2.
CC InterPro: IPR001223; Glyco_hydro_18.
CC Pfam: PF00704; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
CC Glycoprotein; Fertilization; Signal.
CC SIGNAL 1 21 BY SIMILARITY.
CC CHAIN 22 539 OVIDUCT-SPECIFIC GLYCOPROTEIN.
CC CARBOHYD 402 402 N-LINKED (GLCNAC... ) (POTENTIAL).
CC CONFLICT 13 13 M -> V (IN REF. 2).
CC CONFLICT 122 122 K -> N (IN REF. 2).
CC CONFLICT 282 282 A -> V (IN REF. 2).

Query Match 66.7%; Score 64; DB 1; Length 537;
Best Local Similarity 72.2%; Pred. No. 0.01;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTLKNRNPNTLLSVGG 19
DB 82 NQKERNRGLKTLISVG 99

RESULT 7
OGP_BOVIN STANDARD; PRT; 537 AA.
AC Q28042;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN) (FRAGMENT).
GN OVGPI OR OGP.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-47.
RC TISSUE=oviduct;
RX MEDLINE=94257768; PubMed=8199272;
RA Sendai Y., Abe H., Kikuchi M., Satoh T., Hoshi H.;
RA "Purification and molecular cloning of bovine oviduct-specific
RT glycoprotein.";
RT Biol. Reprod. 50:927-934(1994).
CC -1- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -1- TISSUE SPECIFICITY: OVIDUCT.
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D16639; BAA04065.1; -.
CC InterPro: IPR001579; Chitinase_2.
CC InterPro: IPR001223; Glyco_hydro_18.
CC Pfam: PF00704; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
CC Glycoprotein; Fertilization; Signal.
CC NON_TER 1 1
CC SIGNAL <1 18
CC CHAIN 19 537 OVIDUCT-SPECIFIC GLYCOPROTEIN.
CC CARBOHYD 399 399 N-LINKED (GLCNAC... ) (POTENTIAL).
CC SEQUENCE 537 AA; 59617 MW; CFCDFE6F0212D791 CRC64;

Query Match 66.7%; Score 64; DB 1; Length 537;
Best Local Similarity 72.2%; Pred. No. 0.01;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTLKNRNPNTLLSVGG 19
DB 79 NQKERNRGLKTLISVG 96

```

```
RESULT 8
OGP_PAPAN STANDARD; PRT; 623 AA.
ID OGP_PAPAN
AC P36718;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN).
GN OVGPI OR OGP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
OC Cercopitheinae; Papio.
OX NCBI_TaxID=9535;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=98244335; PubMed=9584944;
RA Verhage H.G., Fazleabas A.T., Mavrogianis P.A., O'Day-Bowman M.B.,
RA Donnelly K.M., Arias E.B., Jaffe R.C.;
RT "The baboon oviduct: Characteristics of an oestradiol-dependent
RT oviduct-specific glycoprotein."
RL Hum. Reprod. Update 3:541-552(1997).
RN [2]
RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=91367180; PubMed=1716345;
RA Donnelly K.M., Fazleabas A.T., Verhage H.G., Mavrogianis P.A.,
RA Jaffe R.C.;
RT "Cloning of a recombinant complementary DNA to a baboon (Papio
RT anubis) estradiol-dependent oviduct-specific glycoprotein."
RL Mol. Endocrinol. 5:356-364(1991).
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- DEVELOPMENTAL STAGE: AT THE TIME OF OVULATION.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U09550; AAB86946.1; -
DR EMBL; U58010; AAB04126.1; -
DR EMBL; U58001; AAB04126.1; JOINED.
DR EMBL; U58002; AAB04126.1; JOINED.
DR EMBL; U58003; AAB04126.1; JOINED.
DR EMBL; U58004; AAB04126.1; JOINED.
DR EMBL; U58005; AAB04126.1; JOINED.
DR EMBL; U58006; AAB04126.1; JOINED.
DR EMBL; U58007; AAB04126.1; JOINED.
DR EMBL; U58008; AAB04126.1; JOINED.
DR EMBL; U58009; AAB04126.1; JOINED.
DR MIM; 603578; -
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 678 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 477 477 M -> T (IN REF. 2).
FT CONFLICT 511 511 S -> P (IN REF. 2).
FT CONFLICT 514 514 Y -> H (IN REF. 2).
FT CONFLICT 676 676 E -> Q (IN REF. 2).
SQ SEQUENCE 678 AA; 75421 MW; 245F2CDE92768B CRC64;

Query Match 66.7%; Score 64; DB 1; Length 678;
Best Local Similarity 72.2%; Pred. No. 0.013;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTLKRNPNLKTLLSVGG 19
DB 82 NKLKRNRELKTLISGG 99

RESULT 9
OGP_PAPAN STANDARD; PRT; 678 AA.
ID OGP_PAPAN
AC Q12889; Q15841;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN).
GN OVGPI OR OGP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=95112556; PubMed=7819450;
RA Arias E.B., Verhage H.G., Jaffe R.C.;
RT "Complementary deoxyribonucleic acid cloning and molecular
RT characterization of an estrogen-dependent human oviductal
RT glycoprotein."
RL Biol. Reprod. 51:685-694(1994).
RN [2]
RP SEQUENCE FROM N.A.
RA Jaffe R.C.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBSJ databases.
CC -!- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -!- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -!- TISSUE SPECIFICITY: OVIDUCT.
CC -!- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; U09550; AAB86946.1; -
DR EMBL; U58010; AAB04126.1; -
DR EMBL; U58001; AAB04126.1; JOINED.
DR EMBL; U58002; AAB04126.1; JOINED.
DR EMBL; U58003; AAB04126.1; JOINED.
DR EMBL; U58004; AAB04126.1; JOINED.
DR EMBL; U58005; AAB04126.1; JOINED.
DR EMBL; U58006; AAB04126.1; JOINED.
DR EMBL; U58007; AAB04126.1; JOINED.
DR EMBL; U58008; AAB04126.1; JOINED.
DR EMBL; U58009; AAB04126.1; JOINED.
DR MIM; 603578; -
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 678 OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT CARBOHYD 402 402 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 580 580 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 596 596 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 648 648 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 477 477 M -> T (IN REF. 2).
FT CONFLICT 511 511 S -> P (IN REF. 2).
FT CONFLICT 514 514 Y -> H (IN REF. 2).
FT CONFLICT 676 676 E -> Q (IN REF. 2).
SQ SEQUENCE 678 AA; 75421 MW; 245F2CDE92768B CRC64;

Query Match 66.7%; Score 64; DB 1; Length 678;
Best Local Similarity 72.2%; Pred. No. 0.013;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
```

Tue Dec 11 08:46:47 2001

```

Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTLKRNPNKLTLLSVGG 19
   | | | | | | | | | |
Db 82 NKLKERNRELKLTLLSIGG 99

RESULT 10
OGP_MOUSE STANDARD; PRT: 721 AA.
AC 062010;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN) (ZP-0).
GN OVGPI OR OGP OR CHIT5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR; TISSUE=Oviduct;
RX MEDLINE=96115001; PubMed=7492680;
RA Sendai Y., Komiya H., Suzuki K., Onuma T., Kikuchi M., Hoshi H.,
RA Araki Y.;
RT "Molecular cloning and characterization of a mouse oviduct-specific
glycoprotein.";
RL Biol. Reprod. 53:285-294(1995).
CC -1- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -1- TISSUE SPECIFICITY: EPITHELIAL CELLS OF THE OVIDUCT.
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D32137; BAA06863.1; -
CC MGD; MGI:106661; Chit5.
CC InterPro; IPR001579; Chitinase_2.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Repeat; Signal.
FT SIGNAL 1 21
FT CHAIN 22 721
FT DOMAIN 486 632
FT
FT CARBOHYD 402 402
FT CARBOHYD 442 442
FT CARBOHYD 469 469
FT SEQUENCE 721 AA; 78807 MW; 37246C8F01665652 CRC64;

Query Match 66.7%; Score 64; DB 1; Length 721;
Best Local Similarity 72.2%; Pred. No. 0.014;
Matches 13; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 NTLKRNPNKLTLLSVGG 19
   | | | | | | | | | |
Db 82 NKLKERNRELKLTLLSIGG 99

RESULT 11
OGP_MESAU STANDARD; PRT: 671 AA.
ID OGP_MOUSE

```

```

AC Q60557; Q60526;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (OVIDUCTAL GLYCOPROTEIN)
DE (OVIDUCTIN) (ESTROGEN-DEPENDENT OVIDUCT PROTEIN) (ZP-0).
GN OVGPI OR OGP.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Oviduct;
RX MEDLINE=96115007; PubMed=7492686;
RA Suzuki K., Sendai Y., Onuma T., Hoshi H., Hiroi M., Araki Y.;
RT "Molecular characterization of a hamster oviduct-specific
glycoprotein.";
RL Biol. Reprod. 53:345-354(1995).
RN [2]
RP SEQUENCE OF 22-671 FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=96192955; PubMed=8607967;
RA Paquette Y., Merlen Y., Malette B., Bleau G.;
RT "Allelic polymorphism in the hamster oviductin gene is due to a
variable number of mucin-like tandem repeats.";
RL Mol. Reprod. Dev. 42:388-396(1995).
RN [3]
RP SEQUENCE OF 14-671 FROM N.A., AND REVISIONS.
RC TISSUE=Oviduct;
RA Paquette Y.;
RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 22-39.
RX MEDLINE=94058981; PubMed=8240241;
RA Malette B., Bleau G.;
RT "Biochemical characterization of hamster oviductin as a sulphated
zona pellucida-binding glycoprotein.";
RL Biochem. J. 293:437-445(1993).
CC -1- FUNCTION: BINDS TO OOCYTE ZONA PELLUCIDA IN VIVO. MAY PLAY A ROLE
CC IN THE FERTILIZATION PROCESS AND/OR EARLY EMBRYONIC DEVELOPMENT.
CC MIGHT ACT AS A PROTECTIVE SECRETION INFLUENCING THE FIRST STEPS OF
CC THE REPRODUCTIVE PROCESS NECESSARY FOR THE NORMAL TRIGGERING OF
CC FERTILIZATION AND EARLY EMBRYONIC DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: SECRETORY GRANULES.
CC -1- TISSUE SPECIFICITY: OVIDUCT.
CC -1- PTM: HIGHLY O-GLYCOSYLATED AND ALSO N-GLYCOSYLATED.
CC -1- SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; D32218; BAA06977.1; -
CC EMBL; U15048; AAC53584.1; -
CC InterPro; IPR001579; Chitinase_2.
CC InterPro; IPR001223; Glyco_hydro_18.
CC Pfam; PF00704; Glyco_hydro_18; 1.
CC PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Fertilization; Repeat; Signal.
FT SIGNAL 1 21
FT CHAIN 22 671
FT DOMAIN 490 609
FT REPEAT 490 504 1.
FT REPEAT 505 519 2.
FT REPEAT 520 534 3.
FT REPEAT 535 549 4.
FT REPEAT 550 564 5.
FT
FT OVIDUCT-SPECIFIC GLYCOPROTEIN.
FT 8 X 15 AA TANDEM REPEATS.

```


| | | | | |
|------------|--------------------------------------------------------------------------------------------------------------------------|-------------------|-----------|-------------------------------------|
| FT | REPEAT | 565 | 579 | 6. |
| FT | REPEAT | 580 | 594 | 7. |
| FT | REPEAT | 595 | 609 | 8. |
| FT | CARBOHYD | 402 | 402 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 511 | 511 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 526 | 526 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 541 | 541 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 556 | 556 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 571 | 571 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CARBOHYD | 586 | 586 | N-LINKED (GLCNAC. . .) (POTENTIAL). |
| FT | CONFLICT | 26 | 26 | C -> A (IN REF. 4). |
| FT | CONFLICT | 33 | 33 | H -> I (IN REF. 4). |
| FT | CONFLICT | 137 | 137 | D -> G (IN REF. 2 AND 3). |
| FT | CONFLICT | 153 | 153 | R -> Q (IN REF. 2 AND 3). |
| FT | CONFLICT | 165 | 165 | F -> Y (IN REF. 2 AND 3). |
| FT | CONFLICT | 193 | 193 | O -> L (IN REF. 2 AND 3). |
| FT | CONFLICT | 531 | 545 | MISSING (IN REF. 2 AND 3). |
| FT | CONFLICT | 595 | 595 | T -> I (IN REF. 2 AND 3). |
| FT | SEQUENCE | 671 AA; | 73250 MW; | BB57E0E514EC1972 CRC64; |
| QY | 2 | NTLKNRPNLKTLLSVGG | 19 | |
| DB | 82 | NKLKERNALKTLLSVGG | 99 | |
| RESULT 12 | | | | |
| CHIL_COCIM | | STANDARD; | PRT; | 427 AA. |
| ID | CHIL_COCIM | | | |
| AC | P54196; | | | |
| DT | 01-OCT-1996 (Rel. 34, Created) | | | |
| DT | 01-OCT-1996 (Rel. 34, Last sequence update) | | | |
| DT | 01-OCT-1996 (Rel. 34, Last annotation update) | | | |
| DE | ENDOCHITINASE 1 PRECURSOR (EC 3.2.1.14) (COMPLEMENT-FIXATION ANTIGEN) | | | |
| DE | (CF-ANTIGEN) (CF-AG). | | | |
| DE | CTS1. | | | |
| OS | Coccidioides immitis. | | | |
| GN | Coccidioides immitis. | | | |
| OC | Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes; | | | |
| OC | Ongygenales; mitosporic Onygenales; Coccidioides. | | | |
| NCBI_TaxID | 5501; | | | |
| [1] | SEQUENCE FROM N.A. | | | |
| RN | STRAIN=C735; | | | |
| RC | MEDLINE=96144270; PubMed=8566773; | | | |
| RX | Pishko E.J., Kirkland T.N., Cole G.T.; | | | |
| RA | "Isolation and characterization of two chitinase-encoding genes | | | |
| RT | (cts1, cts2) from the fungus Coccidioides immitis."; | | | |
| RT | Gene 167:173-177(1995). | | | |
| RL | [2] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | STRAIN=SILVEIRA; | | | |
| RA | Yang C., Zhu Y., Magee D.M., Cox R.A.; | | | |
| RL | Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases. | | | |
| CC | -1- CATALYTIC ACTIVITY: HYDROLYSIS OF THE 1,4-BETA-LINKAGES OF | | | |
| CC | N-ACETYL-D-GLUCOSAMINE POLYMERS OF CHITIN. | | | |
| CC | -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL | | | |
| CC | HYDROLASES). | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration | | | |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - | | | |
| CC | the European Bioinformatics Institute. There are no restrictions on its | | | |
| CC | use by non-profit institutions as long as its content is in no way | | | |
| CC | modified and this statement is not removed. Usage by and for commercial | | | |
| CC | entities requires a license agreement (See http://www.isb-sib.ch/announce/ | | | |
| CC | or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |
| DR | EMBL; L41663; AAA92643.1; - | | | |
| DR | EMBL; U51271; AAA96515.1; - | | | |
| DR | EMBL; U33265; AAB06687.1; - | | | |

```

DR HSSP; P07254; ICTN.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Hydrolase; glycosidase; Chitin degradation; Chitin-binding; Signal;
KW Glycoprotein.
DR SIGNAL 1 ? POTENTIAL.
DR CHAIN ? 427 ENDOCHITINASE 1.
FT CARBOHYD 387 387 N-LINKED (GLCNAC... ) (POTENTIAL).
FT CONFLICT 15 47 RWLSRLCLFCELGREMETLSTVAVTVVTDIO -> VQAS
FT CONFLICT 199 199 SMSMPNYYVPVPEAPGEGPSRVYFVNW (IN REF. 2).
FT CONFLICT 427 AA; 47629 MW; 1C396DBDB1A7001A CRC64;
SQ SEQUENCE 427 AA; 47629 MW; 1C396DBDB1A7001A CRC64;

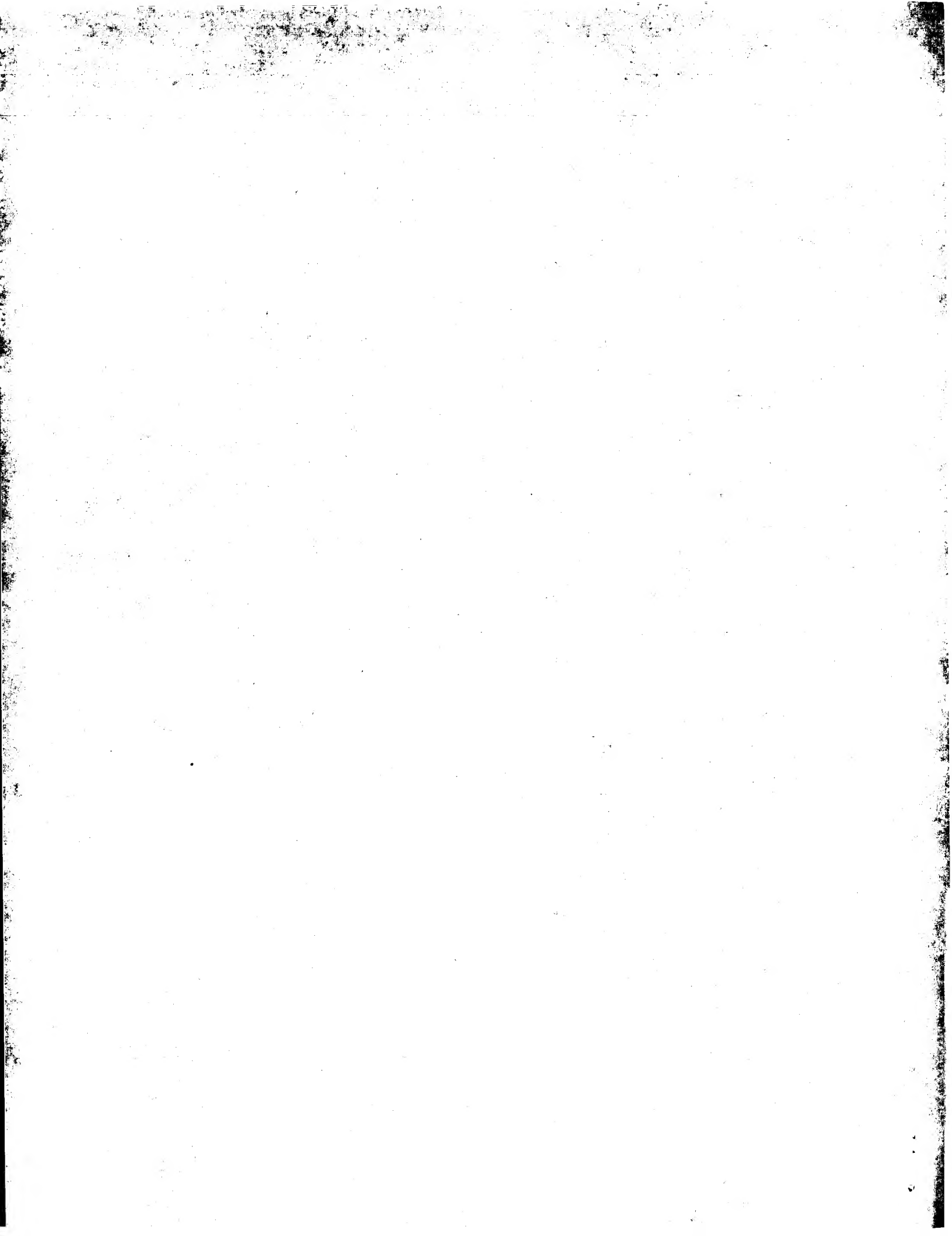
Query Match 62.5%; Score 60; DB 1; Length 427;
Best Local Similarity 75.0%; Pred. No. 0.035;
Matches 12; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 4 LKNRPNLKTLLSVGG 19
|| | |||||:|
Db 115 LKNNRNLTLLSIGG 130

RESULT 13
PARA_MYCPN STANDARD; PRT; 270 AA.
ID PARA_MYCPN AC Q50314;
DI 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PARA FAMILY PROTEIN MPN688.
DN MPN688 OR MP154.
OS Mycoplasma pneumoniae.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2104;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=96177562; PubMed=8604303;
RA Hilbert H., Himmelreich R., Plagens H., Herrmann R.;
RT "Sequence analysis of 56 kb from the genome of the bacterium
RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a
RT cluster of ribosomal protein genes.";
RL Nucleic Acids Res. 24:628-639(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29342 / M129;
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [7]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [8]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [9]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [10]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [11]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [12]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [13]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [14]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [15]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [16]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [17]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [18]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [19]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [20]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [21]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [22]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [23]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [24]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [25]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [26]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [27]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [28]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996).
RN [29]
RP SEQUENCE FROM N.A.
RX MEDLINE=97105885; PubMed=8948633;
RA Himmelreich R., Hilbert H., Plagens H., Pirkel E., Li B.-C.,
RA Herrmann R.;
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma
RT pneumoniae.";
RL Nucleic Acids Res. 24:4420-4449(1996
```

1 LNTLKNRPNLKTLLSVGG 19

Search completed: December 7, 2001, 00:39:47
Job time: 901 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run On: December 7, 2001, 00:38:25 ; Search time 135.48 Seconds
(without alignments)
20.514 Million cell updates/sec

Title: US-09-164-862B-2

Perfect score: 96
Sequence: 1 LNTLKRNPNLKTLLSVGG 19

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

```
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

```
Database :
SPTRMBL_17.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
```

SUMMARIES

| Result No. | Score | Query | | Length | DB | ID | Description |
|------------|-------|-------|-----|--------|--------|--------------------|-------------|
| | | Match | % | | | | |
| 1 | 96 | 100.0 | 332 | 6 | O18949 | O18949 bos taurus | |
| 2 | 96 | 100.0 | 383 | 6 | Q29411 | Q29411 sus scrofa | |
| 3 | 84 | 87.5 | 352 | 11 | Q9WTV1 | Q9WTV1 rattus norv | |
| 4 | 76 | 79.2 | 381 | 11 | Q99J84 | Q99J84 mus musculu | |
| 5 | 74 | 77.1 | 396 | 11 | Q9D7Q1 | Q9D7Q1 mus musculu | |
| 6 | 72 | 75.0 | 699 | 2 | Q48494 | Q48494 kurtzia zop | |
| 7 | 71 | 74.0 | 717 | 2 | Q9KHB3 | Q9KHB3 bacillus ci | |
| 8 | 68 | 70.8 | 262 | 11 | Q9D7W6 | Q9D7W6 mus musculu | |
| 9 | 68 | 70.8 | 472 | 11 | Q9JLN1 | Q9JLN1 mus musculu | |
| 10 | 68 | 70.8 | 473 | 11 | Q9D803 | Q9D803 mus musculu | |
| 11 | 68 | 70.8 | 473 | 11 | Q99PH2 | Q99PH2 mus musculu | |
| 12 | 67 | 69.8 | 599 | 2 | Q9KED7 | Q9KED7 bacillus ha | |
| 13 | 66 | 68.8 | 491 | 2 | P94289 | P94289 bacillus ci | |
| 14 | 64 | 66.7 | 120 | 11 | Q99MT5 | Q99MT5 mus musculu | |
| 15 | 64 | 66.7 | 126 | 6 | Q9XT30 | Q9XT30 macaca radi | |
| 16 | 64 | 66.7 | 332 | 10 | O81854 | O81854 arabidopsi | |
| 17 | 64 | 66.7 | 366 | 10 | O81863 | O81863 arabidopsi | |
| 18 | 64 | 66.7 | 398 | 10 | O81862 | O81862 arabidopsi | |
| 19 | 64 | 66.7 | 399 | 10 | O81861 | O81861 arabidopsi | |

ALIGNMENTS

```

RESULT 1
ID O18949 PRELIMINARY; PRT; 332 AA.
AC O18949;
DT 01-JAN-1998 (T-EMBLrel. 05, Created)
DT 01-JAN-1998 (T-EMBLrel. 05, Last sequence update)
DT 01-JUN-2001 (T-EMBLrel. 17, Last annotation update)
DE CHITINASE-LIKE PROTEIN 1 (FRAGMENT).
GN CLP-1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
[1]
RN SEQUENCE FROM N.A.
RA Recklies A.D., White C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF011373; AAB64304.1; -.
DR InterPro: IPR001223; Glyco_hydro_18.
DR Pfam: PF00704; Glyco_hydro_18; 1.
FT non_ter 1
FT non_ter 332
FT non_ter 332
SQ SEQUENCE 332 AA; 37434 MW; 03F1633233486408F CRC64;

```

DE 38 KDA HEPARIN-BINDING GLYCOPROTEIN.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SMOOTH MUSCLE;
 RX MEDLINE=95286589; PubMed=7768902;
 RA Shackleton L.M., Mann D.M., Millis A.J.;
 RT "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
 RT differentiating vascular smooth muscle cells as a member of a group of
 RT proteins associated with tissue remodeling."
 RL J. Biol. Chem. 270:13076-13083(1995).
 [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE-SMOOTH MUSCLE;
 RA Shackleton L.M., Mann D.M., Millis A.J.T.;
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 DR EMBL; U19900; AAA86482.1; -;
 DR EMBL; Z47803; CAA87764.1; -;
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 KW Heparin-binding.
 SQ SEQUENCE 383 AA; 42443 MW; 3D1039F49910BDC4 CRC64;

Query Match 100.0%; Score 96; DB 6; Length 383;
 Best Local Similarity 100.0%; Pred. No. 3.1e-07;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSSVGG 19
 |||||
 DB 80 LNTLKNRPNLKTLSSVGG 98

RESULT 3
 Q9WTV1 PRELIMINARY; PRT; 352 AA.
 ID Q9WTV1
 AC Q9WTV1;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE GLYCOPROTEIN-39 PRECURSOR (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-LEWIS;
 RA Wendling U., Boots A.M.H., van Eden W.;
 RT "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
 RT potential autoantigen in arthritis."
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF062038; AAD22610.1; -;
 DR HSP; P07254; ICTN.
 DR InterPro; IPR000677; 2S_Globulin.
 DR InterPro; IPR001223; Glyco_hydro_18.
 DR Pfam; PF00704; Glyco_hydro_18; 1.
 DR PRINTS; PR00551; 2SGLOBULIN.
 FT NON_TER 1
 FT NON_TER 352
 SQ SEQUENCE 352 AA; 39391 MW; CBDE991610AC936C CRC64;

Query Match 87.5%; Score 84; DB 11; Length 352;
 Best Local Similarity 89.5%; Pred. No. 2.3e-05;
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSSVGG 19
 |||||
 DB 59 LNTLKNRPNLKTLSSVGG 77

RESULT 4
 Q99J84 PRELIMINARY; PRT; 381 AA.
 ID Q99J84
 AC Q99J84;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE SIMILAR TO CHITINASE 3-LIKE 1 (CARTILAGE GLYCOPROTEIN-39).
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC005611; AAH05611.1; -;
 DR EMBL; BC003780; AAH03780.1; -;
 DR EMBL; BC004734; AAH04734.1; -;
 SQ SEQUENCE 381 AA; 42979 MW; EF6588C5AE9D4450 CRC64;

Query Match 79.2%; Score 76; DB 11; Length 381;
 Best Local Similarity 84.2%; Pred. No. 0.00046;
 Matches 16; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 LNTLKNRPNLKTLSSVGG 19
 |||||
 DB 81 LNTLKNRPNLKTLSSVGG 99

RESULT 5
 Q9D7Q1 PRELIMINARY; PRT; 396 AA.
 ID Q9D7Q1
 AC Q9D7Q1;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE 2300002L19RIK PROTEIN.
 GN 2300002L19RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=TONGUE;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Kuehl P., Lewis S., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Fleischmann W., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Stauber F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
RA Hayashizaki Y.,
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; AK009012; BAB26025.1; -
DR MGD; MGI:1919134; 2300002L19Rik.
DR InterPro; IPR000677; 2S_Globulin.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00551; 2SGLOBULIN.
DR PROSITE; PS01095; CHITINASE_18; 1.
KW Glycosidase; Hydrolase.
SQ SEQUENCE 396 AA; 43413 MW; 32BB674B61AF7123 CRC64;

Query Match 77.1%; Score 74; DB 11; Length 396;
Best Local Similarity 73.7%; Pred. No. 0.001; 2; Indels 0; Gaps 0;
Matches 14; Conservative 3; Mismatches 2;
QY 1 LNTLKNRPNLKTLISVGG 19
DB 80 LNSLTKPNKLLAVGS 98

RESULT 6
Q48494
ID Q48494 PRELIMINARY; PRT; 699 AA.
AC Q48494;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CHITINASE.
OS Kurthia zopfi.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Kurthia.
OX NCBI_TaxID=1650;
RN [1]
RP SEQUENCE FROM N.A.
RA Ikeda S., Toyoda H., Matsuda Y., Ouchi S.;
RT "DNA sequence determination of a chitinase gene chISHI cloned from
RT gram-positive bacterium Kurthia zopfi and its application to
RT biological control of Powdry Mildew of Barley.*;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
CC HYDROLASES).
DR EMBL; D63702; BAA09831.1; -
DR HSSP; P07254; ICTN
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR003610; Chitin_bind3.
DR InterPro; IPR003962; FnIII_repeat.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR001223; Glyco_hydro_18.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00014; FNTPPEIII.
DR SMART; SM00495; ChtBD3; 1.
DR SMART; SM00060; FN3; 2.
DR PROSITE; PS01095; CHITINASE_18; 1.
DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
KW Glycosidase; Hydrolase; Repeat.
SQ SEQUENCE 699 AA; 73494 MW; 2AE9599A604BF513 CRC64;

Query Match 75.0%; Score 72; DB 2; Length 699;
Best Local Similarity 73.7%; Pred. No. 0.0037; 3; Indels 0; Gaps 0;
Matches 14; Conservative 2; Mismatches 3;
QY 1 LNTLKNRPNLKTLISVGG 19

DB 145 LNKLAQINPNLKIISVGG 163

RESULT 7
Q9KHB3
ID Q9KHB3 PRELIMINARY; PRT; 717 AA.
AC Q9KHB3;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE CHITINASE.
GN CH1.
OS Bacillus circulans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=1397;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-2;
RA Wang Y.L., Wang H.Y., Qin M., Zhang Y.Z.;
RT "Nucleotide sequence of a chitinase gene (chil) from Bacillus
RT circulans C-2.*;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF265220; AAF74782.1; -
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR003610; Chitin_bind3.
DR InterPro; IPR003962; FnIII_repeat.
DR InterPro; IPR003961; FN.III.
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00041; fn3; 2.
DR Pfam; PF00704; Glyco_hydro_18; 1.
DR PRINTS; PR00014; FNTPPEIII.
DR SMART; SM00495; ChtBD3; 1.
DR SMART; SM00060; FN3; 2.
DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_1.
KW Hydrolase; Repeat.
SQ SEQUENCE 717 AA; 76118 MW; BB0B019CDE72C198 CRC64;

Query Match 74.0%; Score 71; DB 2; Length 717;
Best Local Similarity 68.4%; Pred. No. 0.0054;
Matches 13; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 LNTLKNRPNLKTLISVGG 19
DB 145 LNKLAQINPNLKIISVGG 163

RESULT 8
Q9D7W6
ID Q9D7W6 PRELIMINARY; PRT; 262 AA.
AC Q9D7W6;
DT 01-JUN-2001 (TRENBLrel. 17, Created)
DT 01-JUN-2001 (TRENBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE 2200003E03RIK PROTEIN.
GN 2200003E03RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=STOMACH;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

Tue Dec 11 08:46:47 2001

us-09-164-862b-2.rspt

0;

0; Gaps

3; Indels

2; Mismatches

13; Conservative

Matches

0;

0; Gaps

3; Indels

2; Mismatches

13; Conservative

Matches

0;

0; Gaps

RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., Fujita M., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).
 CC EMBL: AK008757; BAB25878.1; -.
 DR MGD: MGI:1913535; 2200003E03RIK.
 DR InterPro: IPR001579; Chitinase_2.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR PROSITE: PS01095; CHITINASE_18; 1.
 KW Glycosidase; Hydrolase.
 SQ SEQUENCE 262 AA; 29458 MW; 954AA0F0E1C9851D CRC64;

Query Match 70.8%; Score 68; DB 11; Length 262;
 Best Local Similarity 72.2%; Pred. No. 0.006; 3; Indels 0; Gaps 0;
 Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 NTLKRNPNLKTLLSVGG 19
 | ||||| |||||:|
 Db 21 NDLKRNRSKLTLLAIGG 38

RESULT 9

ID Q9JLNI PRELIMINARY; PRT; 472 AA.
 AC Q9JLNI
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE PUTATIVE CHITINASE PRECURSOR (FRAGMENT).
 GN 2200003E03RIK OR YNL.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AVCC1476668; TISSUE=SKIN;
 RA Price P.A., Harris S.C., Williamson M.K.;
 RT "YNL, A Putative Mouse Chitinase."
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).
 CC EMBL: AF154571; AAF31644.1; -.
 DR MGD: MGI:1913535; 2200003E03RIK.
 DR InterPro: IPR001579; Chitinase_2.
 DR InterPro: IPR002557; Chitin_binding.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam: PF01607; Chitin_bind_2; 1.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR SMART: SM00494; ChtBD2; 1.
 DR PROSITE: PS01095; CHITINASE_18; 1.
 KW Glycosidase; Hydrolase; Signal.
 FT SIGNAL <1 20 POTENTIAL.
 FT CHAIN 21 472 PUTATIVE CHITINASE.
 FT CHAIN 21 472
 SQ SEQUENCE 472 AA; 51872 MW; FF59088512C8A7F0 CRC64;

Query Match 70.8%; Score 68; DB 11; Length 472;
 Best Local Similarity 72.2%; Pred. No. 0.011;

0;

0; Gaps

3; Indels

2; Mismatches

13; Conservative

Matches

0;

0; Gaps

3; Indels

2; Mismatches

13; Conservative

Matches

0;

0; Gaps

QY 2 NTLKRNPNLKTLLSVGG 19
 | ||||| |||||:|

Db 80 NDLKRNRSKLTLLAIGG 97
 | ||||| |||||:|

RESULT 10

Q9D803
 ID Q9D803 PRELIMINARY; PRT; 473 AA.
 AC Q9D803;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE 2200003E03RIK PROTEIN.
 GN 2200003E03RIK
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=STOMACH;
 RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gofobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., Fujita M., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,
 RA Hayashizaki Y.,
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).

CC -!- SIMILARITY: BELONGS TO CHITINASE CLASS II (FAMILY 18 OF GLYCOSYL
 CC HYDROLASES).
 CC EMBL: AK008633; BAB25795.1; -.
 DR MGD: MGI:1913535; 2200003E03RIK.
 DR InterPro: IPR001579; Chitinase_2.
 DR InterPro: IPR002557; Chitin_binding.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam: PF01607; Chitin_bind_2; 1.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR SMART: SM00494; ChtBD2; 1.
 DR PROSITE: PS01095; CHITINASE_18; 1.
 KW Glycosidase; Hydrolase.
 SQ SEQUENCE 473 AA; 52003 MW; 333C87447476695 CRC64;

Query Match 70.8%; Score 68; DB 11; Length 473;
 Best Local Similarity 72.2%; Pred. No. 0.011;

Matches 13; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 NTLKRNPNLKTLLSVGG 19
 | ||||| |||||:|

Db 81 NDLKRNRSKLTLLAIGG 98
 | ||||| |||||:|

RESULT 11

Q99PH2
 ID Q99PH2 PRELIMINARY; PRT; 473 AA.
 AC Q99PH2;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)

Q99MJ5 ID Q99MJ5 PRELIMINARY; PRT; 120 AA.
AC Q99MJ5;
AD 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE OVIDUCTAL GLYCOPROTEIN (FRAGMENT).
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS: TISSUE=OVIDUCT;
RA Bhatt P., Woodbirdi S., Natraj U.:
RT "Molecular cloning and overexpression of murine oviductal glycoprotein
RT and characterization of the protein."
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJP databases.
DR EMBL: AF334669; AAK32144.1; -.

```
FT  NON_TER      1      1
FT  NON_TER     120     120
SQ  SEQUENCE    120 AA; 13770 MW; 4FC25603E58A7877 CRC64;

Query Match      66.7%; Score 64; DB 11; Length 120;
Best Local Similarity 72.2%; Pred. No. 0.012;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  2  NTLKRNPNLKTLTSLVGG 19
    | | | | | | | | | | | |
Db  63  NKLKERNRELKTLTSLIGG 80

RESULT 15
Q9XT30
ID  Q9XT30      PRELIMINARY;      PRT; 126 AA.
AC  Q9XT30;
DT  01-NOV-1999 (TREMBLrel. 12, Created)
DT  01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT  01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE  OVIDUCT-SPECIFIC GLYCOPROTEIN PRECURSOR (FRAGMENT).
OS  Macaca radiata (BonneNet monkey).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC  Cercopithecoidea; Macaca.
OX  NCBI_TaxID=9548;
RN  [1]
RN  SEQUENCE FROM N.A.
RP  TISSUE=OVIDUCT;
RA  Natraj U.;
RT  "Partial cloning and sequencing of a cDNA encoding Bonnet monkey
RL  [Macaca radiata] oviduct specific protein.";
RL  Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AF132215; AAD39144.1; -
DR  InterPro; IPR001223; Glyco_hydro_18.
DR  Pfam; PF00704; Glyco_hydro_18; 1.
FT  NON_TER      1      1
FT  NON_TER     126     126
SQ  SEQUENCE    126 AA; 14444 MW; 6E39F05E15815812 CRC64;

Query Match      66.7%; Score 64; DB 6; Length 126;
Best Local Similarity 72.2%; Pred. No. 0.012;
Matches 13; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY  2  NTLKRNPNLKTLTSLVGG 19
    | | | | | | | | | | | |
Db  72  NKLKERNRELKTLTSLIGG 89

Search completed: December 7, 2001, 00:38:26
Job time: 905 sec
```


Tue Dec 11 08:46:48 2001

XX YKL-40 (40 kDa) was purified from human osteosarcoma MG63 cells.
 CC The N-terminal sequence is shown in AAR07045; the full coding region
 CC of the YKL-40 gene is given in AAQ85245. Homology of the N-terminal
 CC and 2 internal peptides (AAR07046-47) with a bacterial polysaccharide
 CC hydrolase suggests that YKL-40 degrades polysaccharide components
 CC of connective tissue. YKL-40 is a marker of e.g. metastatic breast
 CC cancer and inflammatory or degenerative joint diseases.
 XX
 CC Sequence 7 AA;
 SQ

Query Match 100.0%; Score 34; DB 16; Length 7;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
 Db 1 lrlgapa 7
 |||||

RESULT 2
 AAY79493
 ID AAY79493 standard; Peptide; 7 AA.
 AC AAY79493;
 XX
 XX 01-AUG-2000 (first entry)
 DT
 XX Human cancer marker YKL-40 internal peptide B.
 XX
 XX YKL-40; human; lung cancer; bronchus cancer; colorectal cancer;
 KW prostate cancer; breast cancer; pancreas cancer; stomach cancer;
 KW ovary cancer; bladder cancer; brain cancer; oesophagus cancer;
 KW cervix cancer; melanoma; uterine endometrial cancer;
 KW oral cavity cancer; pharynx cancer; liver cancer; kidney cancer;
 KW biliary tract cancer; small bowel cancer; appendix cancer;
 KW salivary gland cancer; thyroid gland cancer; testis cancer;
 KW adrenal gland cancer; osteosarcoma; chondrosarcoma; liposarcoma;
 KW malignant fibrous histiocytoma; infection; pneumonia; meningitis;
 KW arthritis; rheumatoid arthritis; osteoarthritis; fibrosis;
 KW liver cirrhosis; marker; diagnosis; prognosis.
 KW
 XX Homo sapiens.
 OS
 XX
 XX WO200019206-A1.
 PN
 XX
 XX 06-APR-2000.
 PD
 XX 29-SEP-1999; 99WO-US22615.
 PF
 XX 01-OCT-1998; 98US-0164862.
 PR
 XX (REGC) UNIV CALIFORNIA.
 PA
 XX Price PA, Johansen JS;
 PI
 XX WPI; 2000-303485/26.
 DR
 XX Novel methods for detecting cancers and evaluating the prognosis of
 PT cancer using YKL-40 as a marker of cancer -
 PT
 XX Disclosure; Page 102; l1lpp; English.
 PS
 XX This peptide represents internal peptide B of human YKL-40 mature
 CC polypeptide. YKL-40 is a 40 kDa protein having Tyr, Lys and Leu as
 CC its N-terminal residues (hence, YKL-40). The protein can be
 CC obtained from osteosarcoma cell line Mg63. YKL-40 is a mammalian
 CC member of the chitinase family that is suggested to degrade the
 CC polysaccharide components in connective tissue and/or is a lectin
 CC that binds to specific glycan structures in the extracellular
 CC environment of cells. YKL-40 is useful as a marker for the
 CC presence or absence of a cancer and for the prognosis of a cancer.

CC A claimed method for estimating survival length of cancer patients
 CC comprises obtaining a biological sample from the cancer patient and
 CC measuring the level of YKL-40 in the sample, a higher level than in
 CC healthy humans being indicative of reduced survival expectancy.
 CC The sample is obtained from a cancer patient having at least a
 CC preliminary diagnosis of cancer selected from lung, bronchus,
 CC colorectal, prostate, breast, pancreas, stomach, ovary, urinary
 CC bladder, brain, central nervous system, peripheral nervous system,
 CC oesophagus, cervix, melanoma, uterine endometrial, oral cavity,
 CC pharynx, liver, kidney, biliary tract, small bowel, appendix,
 CC salivary gland, thyroid gland, testes, or adrenal gland cancer, or
 CC osteosarcoma, chondrosarcoma, liposarcoma, or malignant fibrous
 CC histiocytoma. Levels of the YKL-40 marker are elevated in
 CC pathologies associated with tissue remodeling, e.g. degenerative
 CC bone diseases such as rheumatoid arthritis, osteoarthritis, fibrosis,
 CC cirrhosis of the liver, and cancer, especially breast, colon,
 CC prostate, or lung cancer. The marker can be used to identify high
 CC risk patients, and so allow selection of appropriate therapeutic
 CC regimens. The methods may also be used to detect bacterial
 CC infections, such as bacterial pneumonia and meningitis, as these
 CC cause an elevation in YKL-40 levels, as well as diseases
 CC characterized by macrophage activation, e.g. giant cell arteritis.
 CC The YKL-40 marker may also be used to evaluate treatment efficacy,
 CC to check for recurrence of a cancer, to monitor terminal phase
 CC patients, and to check the efficacy of surgical removal of a
 CC primary tumor. The methods allow estimation of the survival time
 CC of patients with cancers, especially prostate, lung or colorectal
 CC cancer, where the colorectal cancer is Duke's stage A, B, C, or D.
 XX
 XX Sequence 7 AA;
 SQ

Query Match 100.0%; Score 34; DB 21; Length 7;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
 Db 1 lrlgapa 7
 |||||

RESULT 3
 AAW45171
 ID AAW45171 standard; peptide; 9 AA.
 XX
 XX AAW45171;
 AC
 XX 28-APR-1998 (first entry)
 DT
 XX Human cartilage glycoprotein 39 derived peptide #72.
 DE
 XX Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
 KW immunological tolerance; T-cell; human cartilage glycoprotein 39;
 KW HC gp-39; rheumatoid arthritis; epitope.
 XX
 XX Synthetic.
 OS
 XX Homo sapiens.
 OS
 XX WO9740068-A1.
 PN
 XX 30-OCT-1997.
 PD
 XX 22-APR-1997; 97WO-EP02051.
 PF
 XX 24-APR-1996; 96EP-0201106.
 PR
 XX (ALKU) AKZO NOBEL NV.
 PA
 XX Boots AMH, Verheijden GFM;
 PI
 XX WPI; 1997-535775/49.
 DR
 XX Peptide suitable for use in antigen specific immunosuppressive
 PT

PT therapy - resembles or mimics epitope present on HC gp-39, so
 PT inducing systemic immunological tolerance to rheumatoid arthritis
 CC auto:antigen
 XX
 PS Claim 1; Page 76; 82pp; English.
 XX
 CC The present sequence represents a peptide which resembles or mimics an
 CC epitope present on human cartilage glycoprotein 39 (HC gp-39), an
 CC autoantigen in rheumatoid arthritis. The invention relates to peptides
 CC consisting of 16-55 amino acid residues comprising at least one of the
 CC following 19 sequences: LVCYVTSWS; ELCTHIYS; ILYSFANIS; LKTLSSVGG;
 CC FIKSVPPFL; FDGLDLAWL; LYPGRDRQ; YDIAKISQH; LDFISIMTY; FLSIMTYDF;
 CC FRQEDASP; YAVGYMLRL; MRLGAPAS; LAYEICDF; LRGATVHRT; YIKDRQLAG;
 CC LAGAMVWAL; VWALDLDDF; or LDLDDFQGS. They can be used medically in
 CC of T-cell mediated destruction of articular cartilage in autoimmune
 CC diseases (e.g. rheumatoid arthritis). They can also be used to detect
 CC activated autoreactive T cells in an individual. The peptides have a
 CC specific effect on the autoreactive T cells, thus leaving the other
 CC components of the immune system intact, unlike the non-specific
 CC suppressive effect of immunosuppressive drugs, and do not cause toxic
 CC side effects. The peptides are predominantly recognised by autoreactive
 CC T cells from rheumatoid arthritis patients, but rarely by those from
 CC healthy donors.
 XX
 SQ Sequence 9 AA;

 Query Match 100.0%; Score 34; DB 18; Length 9;
 Best Local Similarity 100.0%; Pred. No. 4.3e+05;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 LRLGAPA 7
 Db | | | | | | |
 2 lrlgapa 8

 RESULT 4
 AAG96034
 ID AAG96034 standard; Peptide; 10 AA.
 XX
 AC AAG96034;
 XX
 DT 18-SEP-2001 (first entry)
 XX
 DE Human complementary peptide, SEQ ID NO: 2228.
 XX
 KW Human; complementary peptide; ligand; drug discovery; drug design.
 XX
 OS Homo sapiens.
 XX
 PN WO200142277-A2.
 XX
 PD 14-JUN-2001.
 XX
 PF 13-DEC-2000; 2000WO-GB04776.
 XX
 PR 13-DEC-1999; 99GB-0029464.
 XX
 PA (PROT-) PROTEOM LTD.
 XX
 PI Roberts GW, Heal JR;
 XX
 DR WPI; 2001-408419/43.
 XX
 CC A set of peptide ligands consisting of specific complementary peptides
 CC to proteins encoded by genes of the human genome, useful in an assay
 CC for screening and identifying of one or more novel peptides which are
 CC drug candidates or pro-drugs -
 XX
 PS Example 4; Page 361; 646pp; English.
 XX
 CC The invention relates to a set of complementary peptide ligands

CC generated from the human genome. The complementary peptides
 CC interact with their relevant target proteins encoded in the human
 CC genome. They can be used as reagents in drug discovery and as lead
 CC ligands to facilitate drug design and development. The present
 CC sequence is a complementary peptide provided in the specification.
 XX
 SQ Sequence 10 AA;

 Query Match 100.0%; Score 34; DB 22; Length 10;
 Best Local Similarity 100.0%; Pred. No. 0.88;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 QY 1 LRLGAPA 7
 Db | | | | | | |
 4 lrlgapa 10

 RESULT 5
 AAW45136
 ID AAW45136 standard; peptide; 16 AA.
 XX
 AC AAW45136;
 XX
 DT 28-APR-1998 (first entry)
 XX
 DE Human cartilage glycoprotein 39 derived peptide #37.
 XX
 KW Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
 KW immunological tolerance; T-cell; human cartilage glycoprotein 39;
 KW HC gp-39; rheumatoid arthritis; epitope.
 XX
 OS Synthetic.
 OS Homo sapiens.
 XX
 PN WO9740068-A1.
 XX
 PD 30-OCT-1997.
 XX
 PF 22-APR-1997; 97WO-EP02051.
 XX
 PR 24-APR-1996; 96EP-0201106.
 XX
 PA (ALKU) AKZO NOBEL NV.
 XX
 PI Boots AMH, Verheijden GFM;
 XX
 DR WPI; 1997-535775/49.
 XX
 PT Peptide suitable for use in antigen specific immunosuppressive
 PT therapy - resembles or mimics epitope present on HC gp-39, so
 PT inducing systemic immunological tolerance to rheumatoid arthritis
 auto:antigen
 XX
 PS Claim 2; Page 76; 82pp; English.
 XX
 CC The present sequence represents a peptide which resembles or mimics an
 CC epitope present on human cartilage glycoprotein 39 (HC gp-39), an
 CC autoantigen in rheumatoid arthritis. The invention relates to peptides
 CC consisting of 16-55 amino acid residues comprising at least one of the
 CC following 19 sequences: LVCYVTSWS; FLCTHIYS; ILYSFANIS; LKTLSSVGG;
 CC FIKSVPPFL; FDGLDLAWL; LYPGRDRQ; YDIAKISQH; LDFISIMTY; FLSIMTYDF;
 CC FRQEDASP; YAVGYMLRL; MRLGAPAS; LAYEICDF; LRGATVHRT; YIKDRQLAG;
 CC LAGAMVWAL; VWALDLDDF; or LDLDDFQGS. They can be used medically in
 CC of T-cell mediated destruction of articular cartilage in autoimmune
 CC diseases (e.g. rheumatoid arthritis). They can also be used to detect
 CC activated autoreactive T cells in an individual. The peptides have a
 CC specific effect on the autoreactive T cells, thus leaving the other
 CC components of the immune system intact, unlike the non-specific
 CC suppressive effect of immunosuppressive drugs, and do not cause toxic
 CC side effects. The peptides are predominantly recognised by autoreactive
 CC T cells from rheumatoid arthritis patients, but rarely by those from

CC healthy donors.
XX
SQ Sequence 16 AA;

Query Match 100.0%; Score 34; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | | | |
Db 8 lrlgapa 14

RESULT 6
AAW45137
ID AAW45137 standard; peptide; 16 AA.
XX
AC AAW45137;

XX 28-APR-1998 (first entry)

XX Human cartilage glycoprotein 39 derived peptide #38.

XX Articular cartilage; immunosuppressive therapy; antigen; autoantigen;
KW immunological tolerance; T-cell; human cartilage glycoprotein 39;
KW HC gp-39; rheumatoid arthritis; epitope.

XX Synthetic.

OS Homo sapiens.

XX W09740068-A1.

XX 30-OCT-1997.

XX 22-APR-1997; 97WO-EP02051.

XX 24-APR-1996; 96EP-0201106.

XX (ALKU) AKZO NOBEL NV.

XX Boots AMH, Verheijden GFM;

XX WPI; 1997-535775/49.

XX Peptide suitable for use in antigen specific immunosuppressive

PT therapy - resembles or mimics epitope present on HC gp-39, so

PT inducing systemic immunological tolerance to rheumatoid arthritis

PT auto:antigen

XX Claim 2; Page 76; 82pp; English.

XX The present sequence represents a peptide which resembles or mimics an

CC epitope present on human cartilage glycoprotein 39 (HC gp-39), an

CC autoantigen in rheumatoid arthritis. The invention relates to peptides

CC consisting of 16-55 amino acid residues comprising at least one of the

CC following 19 sequences: LVCYITSW; FLCYHIYS; ILYSFANIS; LKTLISVGG;

CC FIKSVPPFL; FQGLDLAML; LYPGRDRKQ; YDIAKISQH; LDFISIMTY; FISIMTYDF;

CC FRGQEDASP; YAVGYMLRL; MLRLGAPAS; LAYEICDF; LRGATVHRT; YLKDRQLAG;

CC LAGAMVWAL; VNALDLDDF; or LDLDLDFQGS. They can be used medically in

CC antigen specific immunosuppressive therapy, particularly the treatment

CC of T-cell mediated destruction of articular cartilage in autoimmune

CC diseases (e.g. rheumatoid arthritis). They can also be used to detect

CC activated autoreactive T cells in an individual. The peptides have a

CC specific effect on the autoreactive T cells, thus leaving the other

CC components of the immune system intact, unlike the non-specific

CC suppressive effect of immunosuppressive drugs, and do not cause toxic

CC side effects. The peptides are predominantly recognised by autoreactive

CC T cells from rheumatoid arthritis patients, but rarely by those from

XX healthy donors.

XX Sequence 16 AA;

Query Match 100.0%; Score 34; DB 18; Length 16;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | | | |
Db 2 lrlgapa 8

RESULT 7
AAW26751
ID AAW26751 standard; Protein; 383 AA.
XX
AC AAW26751;

XX 11-MAY-1998 (first entry)

XX Bovine whey protein.

XX Bovine whey protein; human cartilage glycoprotein 39; HC gp-39;
KW autoantigen; antigen; autoimmune disease; rheumatoid arthritis;
KW inflammation; arthritogenic protein; immunotherapy; therapy.

XX Bos taurus.

XX W09740149-A1.

XX 30-OCT-1997.

XX 15-APR-1997; 97WO-EP01903.

XX 18-APR-1996; 96US-0634493.

XX (ALKU) AKZO NOBEL NV.

XX Boots AMH, Bos ES, Verheijden GFM;

XX WPI; 1997-535833/49.

XX N-PSDB; AAT99452.

XX Autoantigen proteins homologous to human cartilage glycoprotein 39 -

PT induce arthritis and provide antigen-specific treatment of articular

PT cartilage destruction in autoimmune diseases e.g. rheumatoid

PT arthritis

XX Claim 11; Page 23-25; 35pp; English.

XX This bovine 39 kDa whey protein sequence was deduced from cDNA

CC clones (see AAT99452) isolated from a cattle mammary gland cDNA

CC library using human cartilage glycoprotein 39 (HC gp-39) cDNA as

CC probe. Bovine whey protein was shown to be arthritogenic, inducing

CC arthritis in animals in the same way as described for HC gp-39 (see

CC AAW26750). Arthritogenic proteins such as bovine whey protein, and

CC which display at least 50% amino acid homology to HC gp-39, are

CC very suitable for inducing systemic tolerance of the immune system

CC to homologous autoantigens and can be used to delay and/or suppress

CC arthritic development in mammals. They can induce specific T-cell

CC tolerance to HC gp-39 in patients suffering from T-cell mediated

CC cartilage destruction, such as rheumatoid arthritis. Arthritogenic

CC proteins are also suitable to induce arthritis in animals,

CC preferably mice, e.g. for use in drug screening.

XX Sequence 383 AA;

Query Match 100.0%; Score 34; DB 18; Length 383;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | | | |
Db 245 lrlgapa 251

RESULT 8
AAB03442
ID AAB03442 standard; Protein; 383 AA.
XX
AC AAB03442;
XX
DT 03-JAN-2001 (first entry)
XX
DE Gp38k protein sequence.
XX
KW Gp38k; chemoattractant; cell migration; wound healing; angiogenesis;
KW cancer; vascular trauma; vascular disease; atherosclerosis; restenosis;
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT Peptide 1..24
FT /label= signal_peptide
FT Region 30..37
FT /note= "peptide antibody"
FT Modified-site 59..61
FT /label= glycosylation_site
FT Binding-site 68..75
FT /label= leucine_zipper
FT Binding-site 82..89
FT /label= leucine_zipper
FT Active-site 131..136
FT /label= chitinase_active_site
FT Binding-site 143..146
FT /label= heparin_binding_site
FT Binding-site 147..154
FT /label= hyaluronic_acid_binding_site
FT Binding-site 262..270
FT /label= hyaluronic_acid_binding_site
FT Binding-site 278..281
FT /label= glycosaminoglycan_binding_site
FT Region 354..357
FT /note= "acidic region"
FT Binding-site 368..376
FT /label= hyaluronic_acid_binding_site
XX
PN WO200034469-A1.
XX
PD 15-JUN-2000.
XX
PF 10-DEC-1999; 99WO-US29262.
XX
PR 11-DEC-1998; 98US-0111856.
XX
PA (UUNY) UNIV NEW YORK STATE RES FOUND.
XX
PI Millis AJT;
XX
DR WPI; 2000-431300/37.
XX
PT Clusterin and gp38k-related peptide capable of altering cell migration
PT useful for treating atherosclerosis, cancer and stenosis following
PT vascular trauma or disease
PS Disclosure; Fig 2; 43pp; English.
XX
CC The present sequence is the protein sequence of gp38k. Gp38k, a
CC chemoattractant, is essential for the migration of vascular smooth muscle
CC cells (VSMC). The gene and protein can, therefore, be used to promote
CC wound healing, angiogenesis and vasculogenesis, in the treatment of
CC stenosis following vascular trauma or disease and to treat
CC atherosclerosis, and antisense sequences can be used to treat cancer, as
CC angiogenesis is vital for tumour survival.
XX
SQ Sequence 383 AA;

Query Match 100.0%; Score 34; DB 21; Length 383;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 1 LRLGAPA 7
Db 245 lrlgapa 251
|||||
RESULT 9
AAG35435
ID AAG35435 standard; Protein; 301 AA.
XX
AC AAG35435;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 43285.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 16-APR-1999; 99US-0128714.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.

PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161320.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 94.1%; Score 32; DB 21; Length 301;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 LRLGAPA 7
:|||||

Db 56 irlgapa 62

RESULT 10
AAG35434
ID AAG35434 standard; Protein; 326 AA.
XX
AC AAG35434;
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 43284.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; corn.
XX
OS Zea mays subsp. mays.
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0123788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 05-MAY-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 18-MAY-1999; 99US-0134370.
PR 19-MAY-1999; 99US-0134768.
PR 20-MAY-1999; 99US-0134941.
PR 21-MAY-1999; 99US-0135124.
PR 24-MAY-1999; 99US-0135353.
PR 25-MAY-1999; 99US-0135629.
PR 27-MAY-1999; 99US-0136021.
PR 28-MAY-1999; 99US-0136392.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.

| | | | | |
|----|--------------|---------------|----|--------------------------------------------------------------------------|
| PR | 18-AUG-1999; | 99US-0149426. | AC | AAG43944; |
| PR | 20-AUG-1999; | 99US-0149722. | XX | 18-OCT-2000 (first entry) |
| PR | 20-AUG-1999; | 99US-0149723. | DT | |
| PR | 20-AUG-1999; | 99US-0149929. | XX | Zea mays protein fragment SEQ ID NO: 54985. |
| PR | 23-AUG-1999; | 99US-0149902. | DE | |
| PR | 23-AUG-1999; | 99US-0149930. | XX | |
| PR | 23-AUG-1999; | 99US-0150566. | KW | Protein identification; signal transduction pathway; metabolic pathway; |
| PR | 26-AUG-1999; | 99US-0150884. | KW | hybridisation assay; genetic mapping; gene expression control; promoter; |
| PR | 27-AUG-1999; | 99US-0151065. | KW | termination sequence; corn. |
| PR | 27-AUG-1999; | 99US-0151066. | XX | |
| PR | 27-AUG-1999; | 99US-0151080. | OS | Zea mays subsp. mays. |
| PR | 30-AUG-1999; | 99US-0151303. | XX | |
| PR | 31-AUG-1999; | 99US-0151438. | PN | EP1033405-A2. |
| PR | 01-SEP-1999; | 99US-0151930. | PD | 06-SEP-2000. |
| PR | 07-SEP-1999; | 99US-0152363. | XX | |
| PR | 10-SEP-1999; | 99US-0153070. | XX | 25-FEB-2000; 2000EP-0301439. |
| PR | 13-SEP-1999; | 99US-0153758. | PF | |
| PR | 15-SEP-1999; | 99US-0154018. | XX | 25-FEB-1999; 99US-0121825. |
| PR | 16-SEP-1999; | 99US-0154039. | PR | 05-MAR-1999; 99US-0123180. |
| PR | 20-SEP-1999; | 99US-0154779. | PR | 09-MAR-1999; 99US-0123548. |
| PR | 22-SEP-1999; | 99US-0155139. | PR | 23-MAR-1999; 99US-0125788. |
| PR | 23-SEP-1999; | 99US-0155486. | PR | 25-MAR-1999; 99US-0126264. |
| PR | 24-SEP-1999; | 99US-0155659. | PR | 29-MAR-1999; 99US-0126785. |
| PR | 28-SEP-1999; | 99US-0156458. | PR | 01-APR-1999; 99US-0127462. |
| PR | 29-SEP-1999; | 99US-0156596. | PR | 06-APR-1999; 99US-0128234. |
| PR | 04-OCT-1999; | 99US-0157117. | PR | 08-APR-1999; 99US-0128714. |
| PR | 05-OCT-1999; | 99US-0157753. | PR | 16-APR-1999; 99US-0129845. |
| PR | 06-OCT-1999; | 99US-0157865. | PR | 19-APR-1999; 99US-0130077. |
| PR | 07-OCT-1999; | 99US-0158029. | PR | 21-APR-1999; 99US-0130449. |
| PR | 08-OCT-1999; | 99US-0158232. | PR | 23-APR-1999; 99US-0130510. |
| PR | 12-OCT-1999; | 99US-0158369. | PR | 23-APR-1999; 99US-0130891. |
| PR | 13-OCT-1999; | 99US-0159293. | PR | 28-APR-1999; 99US-0131449. |
| PR | 13-OCT-1999; | 99US-0159294. | PR | 30-APR-1999; 99US-0132048. |
| PR | 13-OCT-1999; | 99US-0159295. | PR | 30-APR-1999; 99US-0132407. |
| PR | 14-OCT-1999; | 99US-0159329. | PR | 04-MAY-1999; 99US-0132484. |
| PR | 14-OCT-1999; | 99US-0159331. | PR | 05-MAY-1999; 99US-0132485. |
| PR | 14-OCT-1999; | 99US-0159637. | PR | 06-MAY-1999; 99US-0132486. |
| PR | 14-OCT-1999; | 99US-0159638. | PR | 06-MAY-1999; 99US-0132487. |
| PR | 18-OCT-1999; | 99US-0159584. | PR | 07-MAY-1999; 99US-0132863. |
| PR | 21-OCT-1999; | 99US-0160741. | PR | 11-MAY-1999; 99US-0134256. |
| PR | 21-OCT-1999; | 99US-0160767. | PR | 14-MAY-1999; 99US-0134218. |
| PR | 21-OCT-1999; | 99US-0160768. | PR | 14-MAY-1999; 99US-0134219. |
| PR | 21-OCT-1999; | 99US-0160770. | PR | 14-MAY-1999; 99US-0134221. |
| PR | 21-OCT-1999; | 99US-0160814. | PR | 14-MAY-1999; 99US-0134370. |
| PR | 21-OCT-1999; | 99US-0160815. | PR | 18-MAY-1999; 99US-0134768. |
| PR | 22-OCT-1999; | 99US-0160980. | PR | 19-MAY-1999; 99US-0134941. |
| PR | 22-OCT-1999; | 99US-0160981. | PR | 20-MAY-1999; 99US-0135124. |
| PR | 22-OCT-1999; | 99US-0160983. | PR | 21-MAY-1999; 99US-0135353. |
| PR | 25-OCT-1999; | 99US-0161404. | PR | 24-MAY-1999; 99US-0135629. |
| PR | 25-OCT-1999; | 99US-0161405. | PR | 25-MAY-1999; 99US-0136021. |
| PR | 25-OCT-1999; | 99US-0161406. | PR | 27-MAY-1999; 99US-0136392. |
| PR | 26-OCT-1999; | 99US-0161359. | PR | 28-MAY-1999; 99US-0136782. |
| PR | 26-OCT-1999; | 99US-0161360. | PR | 03-JUN-1999; 99US-0137222. |
| PR | 26-OCT-1999; | 99US-0161361. | PR | 10-JUN-1999; 99US-0137528. |
| PR | 28-OCT-1999; | 99US-0161920. | PR | 04-JUN-1999; 99US-0137502. |
| PR | 28-OCT-1999; | 99US-0161992. | PR | 07-JUN-1999; 99US-0137724. |
| PR | 28-OCT-1999; | 99US-0161993. | PR | 08-JUN-1999; 99US-0138094. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 10-JUN-1999; 99US-0138540. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 10-JUN-1999; 99US-0138847. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 14-JUN-1999; 99US-0139119. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 16-JUN-1999; 99US-0139452. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 16-JUN-1999; 99US-0139453. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 17-JUN-1999; 99US-0139492. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139454. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139455. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139456. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139457. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139458. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139459. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139460. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139461. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 18-JUN-1999; 99US-0139462. |

Query Match 94.1%; Score 32; DB 21; Length 326;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
:|||||
Db 81 irlgapa 87

RESULT 11
AAG43944
ID AAG43944 standard; Protein; 359 AA.
XX

PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 24-JUN-1999; 99US-0140823.
PR 28-JUN-1999; 99US-0140991.
PR 29-JUN-1999; 99US-0141287.
PR 30-JUN-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 01-JUL-1999; 99US-0142055.
PR 02-JUL-1999; 99US-0142390.
PR 06-JUL-1999; 99US-0142803.
PR 08-JUL-1999; 99US-0142920.
PR 09-JUL-1999; 99US-0142977.
PR 12-JUL-1999; 99US-0143525.
PR 13-JUL-1999; 99US-0143624.
PR 14-JUL-1999; 99US-0144005.
PR 15-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 16-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.

PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161921.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 94.1%; Score 32; DB 21; Length 359;
Best Local Similarity 85.7%; Pred. No. 72;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
:|||||
Db 230 irlgapa 236

RESULT 12
AAG43943
ID AAG43943 standard; Protein; 361 AA.
XX
AC AAG43943;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 54984.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;

Tue Dec 11 08:46:48 2001

us-09-164-862b-3.rag

KW hybridisation assay; genetic mapping; gene expression control; promoter;
 KW termination sequence; corn.

XX Zea mays subsp. mays.

PN EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143542.
 PR 14-JUL-1999; 99US-0143624.
 PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 22-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 25-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 27-AUG-1999; 99US-0151080.
 PR 30-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.

| | | | | |
|------------------------------------------------------------|--------------|---------------|----|------------------------------|
| PR | 07-SEP-1999; | 99US-0152363. | PD | 06-SEP-2000. |
| PR | 10-SEP-1999; | 99US-0153070. | XX | |
| PR | 13-SEP-1999; | 99US-0153758. | PF | 25-FEB-2000; 2000EP-0301439. |
| PR | 15-SEP-1999; | 99US-0154018. | XX | |
| PR | 16-SEP-1999; | 99US-0154039. | PR | 25-FEB-1999; 99US-0121825. |
| PR | 20-SEP-1999; | 99US-0154779. | PR | 05-MAR-1999; 99US-0123180. |
| PR | 22-SEP-1999; | 99US-0155139. | PR | 09-MAR-1999; 99US-0123548. |
| PR | 23-SEP-1999; | 99US-0155486. | PR | 23-MAR-1999; 99US-0125788. |
| PR | 24-SEP-1999; | 99US-0155659. | PR | 25-MAR-1999; 99US-0126264. |
| PR | 28-SEP-1999; | 99US-0156458. | PR | 29-MAR-1999; 99US-0126785. |
| PR | 29-SEP-1999; | 99US-0156596. | PR | 01-APR-1999; 99US-0127462. |
| PR | 04-OCT-1999; | 99US-0157117. | PR | 06-APR-1999; 99US-0128234. |
| PR | 05-OCT-1999; | 99US-0157753. | PR | 08-APR-1999; 99US-0128714. |
| PR | 06-OCT-1999; | 99US-0157865. | PR | 16-APR-1999; 99US-0129845. |
| PR | 07-OCT-1999; | 99US-0158029. | PR | 19-APR-1999; 99US-0130077. |
| PR | 08-OCT-1999; | 99US-0158232. | PR | 21-APR-1999; 99US-0130449. |
| PR | 12-OCT-1999; | 99US-0158369. | PR | 23-APR-1999; 99US-0130510. |
| PR | 13-OCT-1999; | 99US-0159293. | PR | 23-APR-1999; 99US-0130891. |
| PR | 13-OCT-1999; | 99US-0159294. | PR | 28-APR-1999; 99US-0131449. |
| PR | 13-OCT-1999; | 99US-0159295. | PR | 30-APR-1999; 99US-0132048. |
| PR | 14-OCT-1999; | 99US-0159329. | PR | 30-APR-1999; 99US-0132407. |
| PR | 14-OCT-1999; | 99US-0159330. | PR | 04-MAY-1999; 99US-0132484. |
| PR | 14-OCT-1999; | 99US-0159331. | PR | 05-MAY-1999; 99US-0132485. |
| PR | 14-OCT-1999; | 99US-0159637. | PR | 06-MAY-1999; 99US-0132486. |
| PR | 14-OCT-1999; | 99US-0159638. | PR | 07-MAY-1999; 99US-0132487. |
| PR | 18-OCT-1999; | 99US-0159584. | PR | 07-MAY-1999; 99US-0132863. |
| PR | 21-OCT-1999; | 99US-0160741. | PR | 11-MAY-1999; 99US-0134256. |
| PR | 21-OCT-1999; | 99US-0160767. | PR | 14-MAY-1999; 99US-0134218. |
| PR | 21-OCT-1999; | 99US-0160768. | PR | 14-MAY-1999; 99US-0134219. |
| PR | 21-OCT-1999; | 99US-0160770. | PR | 14-MAY-1999; 99US-0134221. |
| PR | 21-OCT-1999; | 99US-0160814. | PR | 14-MAY-1999; 99US-0134370. |
| PR | 21-OCT-1999; | 99US-0160815. | PR | 18-MAY-1999; 99US-0134768. |
| PR | 22-OCT-1999; | 99US-0160980. | PR | 19-MAY-1999; 99US-0134941. |
| PR | 22-OCT-1999; | 99US-0160981. | PR | 20-MAY-1999; 99US-0135124. |
| PR | 22-OCT-1999; | 99US-0160989. | PR | 21-MAY-1999; 99US-0135353. |
| PR | 25-OCT-1999; | 99US-0161404. | PR | 24-MAY-1999; 99US-0135629. |
| PR | 25-OCT-1999; | 99US-0161405. | PR | 25-MAY-1999; 99US-0136021. |
| PR | 25-OCT-1999; | 99US-0161406. | PR | 27-MAY-1999; 99US-0136392. |
| PR | 26-OCT-1999; | 99US-0161359. | PR | 28-MAY-1999; 99US-0136782. |
| PR | 26-OCT-1999; | 99US-0161360. | PR | 01-JUN-1999; 99US-0137222. |
| PR | 26-OCT-1999; | 99US-0161361. | PR | 03-JUN-1999; 99US-0137528. |
| PR | 28-OCT-1999; | 99US-0161920. | PR | 04-JUN-1999; 99US-0137502. |
| PR | 28-OCT-1999; | 99US-0161992. | PR | 07-JUN-1999; 99US-0137724. |
| PR | 28-OCT-1999; | 99US-0161993. | PR | 08-JUN-1999; 99US-0138094. |
| PR | 29-OCT-1999; | 99US-0162142. | PR | 10-JUN-1999; 99US-0138540. |
| Query Match 94.1%; Score 32; DB 21; Length 361; | | | PR | 10-JUN-1999; 99US-0138847. |
| Best Local Similarity 85.7%; Pred. No. 73; | | | PR | 14-JUN-1999; 99US-0139119. |
| Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0; | | | PR | 16-JUN-1999; 99US-0139452. |
| Qy 1 LRLGAPA 7 | | | PR | 16-JUN-1999; 99US-0139453. |
| Db 232 IrlGAPA 238 | | | PR | 17-JUN-1999; 99US-0139492. |
| | | | PR | 18-JUN-1999; 99US-0139454. |
| | | | PR | 18-JUN-1999; 99US-0139455. |
| | | | PR | 18-JUN-1999; 99US-0139456. |
| | | | PR | 18-JUN-1999; 99US-0139457. |
| | | | PR | 18-JUN-1999; 99US-0139458. |
| | | | PR | 18-JUN-1999; 99US-0139459. |
| | | | PR | 18-JUN-1999; 99US-0139460. |
| | | | PR | 18-JUN-1999; 99US-0139461. |
| | | | PR | 18-JUN-1999; 99US-0139462. |
| | | | PR | 18-JUN-1999; 99US-0139463. |
| | | | PR | 18-JUN-1999; 99US-0139750. |
| | | | PR | 18-JUN-1999; 99US-0139763. |
| | | | PR | 21-JUN-1999; 99US-0139817. |
| | | | PR | 22-JUN-1999; 99US-0139899. |
| | | | PR | 23-JUN-1999; 99US-0140353. |
| | | | PR | 23-JUN-1999; 99US-0140354. |
| | | | PR | 24-JUN-1999; 99US-0140695. |
| | | | PR | 28-JUN-1999; 99US-0140823. |
| | | | PR | 29-JUN-1999; 99US-0140991. |
| | | | PR | 30-JUN-1999; 99US-0141287. |
| | | | PR | 01-JUL-1999; 99US-0141842. |
| | | | PR | 01-JUL-1999; 99US-0142154. |
| | | | PR | 02-JUL-1999; 99US-0142055. |

RESULT 13
AAG43942
ID AAG43942 standard; Protein; 363 AA.
XX AC AAG43942;
XX
DT 18-OCT-2000 (first entry)
XX
DE Zea mays protein fragment SEQ ID NO: 54983.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence; Corn.
XX
OS Zea mays subsp. mays.
XX
PN EP1033405-A2.
XX

XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L;
XX WPI: 2000-465970/40.
DR New corn plant and Arabidopsis thaliana sequence-determined DNA
PT fragments, useful for expressing gene products and for controlling
PT expression of a target gene -
XX Claim 14; Page 574; 673pp; English.
XX The present invention describes polynucleotides, such as complete cDNA
CC sequences and/or sequences of genomic DNA encompassing complete genes,
CC portions of genes, and/or intergenic regions, collectively referred to
CC as sequence-determined DNA fragments (SDFs), from corn plants and
CC Arabidopsis thaliana. The SDFs are promoters, structural genes,
CC untranslated regions (UTRs), or 3' termination sequences. They can be
CC used for expressing a gene product and controlling expression of a
CC target gene, either as a promoter, a structural gene, an UTR or as a
CC 3' termination sequence. They are also useful as tools for genetic
CC mapping, and identification of a particular individual plant or for
CC clustering a group of plants with a common trait. AAA78433 to AAA78630
CC and AAB24605 to AAB25099 represent the specifically claimed
CC polynucleotide sequences and polypeptides encoded by them given in the
CC present invention.
XX SQ Sequence 115 AA;
Query Match 91.2%; Score 31; DB 21; Length 115;
Best Local Similarity 85.7%; Pred. No. 39;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 LRLGAPA 7
DB 51 lklgapa 57
RESULT 15
AAB24945
ID AAB24945 standard; Peptide; 117 AA.
XX AC AAB24945;
DT 27-NOV-2000 (first entry)
XX DE Plant SDF encoded polypeptide sequence SEQ List 1 NO:477.
XX Plant; corn; Arabidopsis thaliana; sequence-determined DNA fragment;
KW SDF; genetic mapping; identification; promoter; structural gene; UTR;
KW untranslated region; expression control.
OS Plant.
XX WO2000040695-A2.
PN 13-JUL-2000.
XX 07-JAN-2000; 2000WO-US00466.
PF 08-JAN-1999; 99US-0115293.
PR (CERE-) CERES INC.
XX Alexandrov N, Brover V, Chen X, Subramanian G, Troukhan ME;
PI Zheng L;
XX WPI: 2000-465970/40.
DR New corn plant and Arabidopsis thaliana sequence-determined DNA
XX fragments, useful for expressing gene products and for controlling
PT expression of a target gene -

XX Claim 14; Page 573; 673pp; English.
PS The present invention describes polynucleotides, such as complete cDNA
XX sequences and/or sequences of genomic DNA encompassing complete genes,
CC portions of genes, and/or intergenic regions, collectively referred to
CC as sequence-determined DNA fragments (SDFs), from corn plants and
CC Arabidopsis thaliana. The SDFs are promoters, structural genes,
CC untranslated regions (UTRs), or 3' termination sequences. They can be
CC used for expressing a gene product and controlling expression of a
CC target gene, either as a promoter, a structural gene, an UTR or as a
CC 3' termination sequence. They are also useful as tools for genetic
CC mapping, and identification of a particular individual plant or for
CC clustering a group of plants with a common trait. AAA78433 to AAA78630
CC and AAB24605 to AAB25099 represent the specifically claimed
CC polynucleotide sequences and polypeptides encoded by them given in the
CC present invention.
XX SQ Sequence 117 AA;
Query Match 91.2%; Score 31; DB 21; Length 117;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 LRLGAPA 7
DB 53 lklgapa 59
Search completed: December 7, 2001, 00:23:16
Job time: 2961 sec

us-09-164-862b-3.rag

Tue Dec 11 08:46:48 2001

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:24:39 ; Search time 71.83 seconds
(without alignments)
2.193 Million cell updates/sec

Title: US-09-164-862B-3

Perfect score: 34

Sequence: 1 LRLGAPA 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_AA:*

1: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.*

2: /cgn2_6/ptodata/2/iaa/5B_COMB.pep.*

3: /cgn2_6/ptodata/2/iaa/6A_COMB.pep.*

4: /cgn2_6/ptodata/2/iaa/6B_COMB.pep.*

5: /cgn2_6/ptodata/2/iaa/PTUS_COMB.pep.*

6: /cgn2_6/ptodata/2/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | % Match | Query Length | DB ID | Description |
|------------|-------|---------|--------------|-------|-------------------|
| 1 | 34 | 100.0 | 7 | 2 | US-08-581-527-3 |
| 2 | 34 | 100.0 | 7 | 5 | PCT-US94-07754-3 |
| 3 | 34 | 100.0 | 9 | 4 | US-09-171-705-72 |
| 4 | 34 | 100.0 | 16 | 4 | US-09-171-705-37 |
| 5 | 34 | 100.0 | 16 | 4 | US-09-171-705-38 |
| 6 | 30 | 88.2 | 56 | 1 | US-08-473-981A-8 |
| 7 | 30 | 88.2 | 56 | 2 | US-08-474-087-8 |
| 8 | 28 | 82.4 | 23 | 1 | US-08-466-265-3 |
| 9 | 28 | 82.4 | 205 | 2 | US-08-775-009-37 |
| 10 | 28 | 82.4 | 300 | 3 | US-09-248-335-36 |
| 11 | 28 | 82.4 | 434 | 1 | US-08-682-193A-2 |
| 12 | 28 | 82.4 | 600 | 2 | US-08-679-405-2 |
| 13 | 28 | 82.4 | 600 | 2 | US-08-842-799-2 |
| 14 | 28 | 82.4 | 600 | 4 | US-09-271-778-2 |
| 15 | 28 | 82.4 | 600 | 5 | PCT-US96-11459-2 |
| 16 | 28 | 82.4 | 609 | 4 | US-09-271-778-1 |
| 17 | 28 | 82.4 | 1418 | 3 | US-08-963-825-20 |
| 18 | 28 | 82.4 | 1418 | 4 | US-09-010-999-1 |
| 19 | 28 | 82.4 | 1442 | 2 | US-08-316-650-12 |
| 20 | 28 | 82.4 | 1442 | 5 | PCT-US95-02251-12 |
| 21 | 28 | 82.4 | 1479 | 3 | US-08-840-062-2 |
| 22 | 27 | 79.4 | 360 | 1 | US-08-205-506A-2 |
| 23 | 27 | 79.4 | 360 | 5 | PCT-US94-02389-2 |
| 24 | 27 | 79.4 | 373 | 4 | US-09-039-198A-14 |
| 25 | 27 | 79.4 | 373 | 4 | US-09-039-198A-15 |
| 26 | 27 | 79.4 | 387 | 2 | US-08-486-839-6 |
| 27 | 27 | 79.4 | 387 | 3 | US-09-151-011-6 |

| | | | | | |
|----|------|------|---|-------------------|--------------------|
| 28 | 79.4 | 387 | 4 | US-09-343-623-6 | Sequence 6, Appli |
| 29 | 79.4 | 395 | 3 | US-08-869-696-3 | Sequence 3, Appli |
| 30 | 79.4 | 459 | 4 | US-09-071-709-1 | Sequence 1, Appli |
| 31 | 79.4 | 466 | 2 | US-08-486-839-4 | Sequence 4, Appli |
| 32 | 79.4 | 466 | 3 | US-09-151-011-4 | Sequence 4, Appli |
| 33 | 79.4 | 466 | 4 | US-09-039-198A-2 | Sequence 2, Appli |
| 34 | 79.4 | 466 | 4 | US-09-039-198A-4 | Sequence 4, Appli |
| 35 | 79.4 | 466 | 4 | US-09-343-623-4 | Sequence 4, Appli |
| 36 | 79.4 | 497 | 3 | US-08-869-696-9 | Sequence 9, Appli |
| 37 | 79.4 | 556 | 3 | US-08-869-696-2 | Sequence 2, Appli |
| 38 | 79.4 | 559 | 1 | US-08-368-071-12 | Sequence 12, Appli |
| 39 | 79.4 | 559 | 1 | US-08-458-181-12 | Sequence 12, Appli |
| 40 | 79.4 | 559 | 5 | PCT-US93-02172-12 | Sequence 12, Appli |
| 41 | 79.4 | 2152 | 4 | US-09-036-987A-3 | Sequence 3, Appli |
| 42 | 79.4 | 2152 | 4 | US-09-370-700-3 | Sequence 3, Appli |
| 43 | 76.5 | 99 | 3 | US-09-173-581-8 | Sequence 8, Appli |
| 44 | 76.5 | 99 | 4 | US-09-420-915-8 | Sequence 8, Appli |
| 45 | 76.5 | 168 | 1 | US-07-661-610C-6 | Sequence 6, Appli |

ALIGNMENTS

RESULT 1

US-08-581-527-3

; Sequence 3, Application US/08581527

; Patent No. 5935798

; GENERAL INFORMATION:

; APPLICANT: Price, Paul A.

; APPLICANT: Johansen, Julia S.

; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER

; TITLE OF INVENTION: FOR DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson, P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/581,527

; FILING DATE: 17-APR-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO95/01995

; FILING DATE: 19-JAN-1995

; APPLICATION NUMBER: 08/089,989

; FILING DATE: 09-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Freeman, John W.

; REGISTRATION NUMBER: 29,066

; REFERENCE/DOCKET NUMBER: 07341/011001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-542-5070

; TELEFAX: 617-542-8906

; INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 7 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: peptide

; IMMEDIATE SOURCE:

; CLONE: YKL-40 Internal Peptide B

; FEATURE:

; NAME/KEY: Peptide

; LOCATION: 1...7

; US-08-581-527-3

us-09-164-862b-3.ra1

Tue Dec 11 08:46:48 2001

Query Match 100.0%; Score 34; DB 2; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | | | |
Db 1 LRLGAPA 7

RESULT 2
PCT-US94-07754-3
; Sequence 3, Application PC/TUS9407754
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07754
; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD 3665
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: YKL-40 INTERNAL PEPTIDE B
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..7
; PCT-US94-07754-3

Query Match 100.0%; Score 34; DB 5; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | | | |
Db 1 LRLGAPA 7

RESULT 3
US-09-171-705-72
; Sequence 72, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
; APPLICANT: BOOTS, ANNA M.H.
; APPLICANT: VERHEIJDEN, GILBERTUS F.M.

Query Match 100.0%; Score 34; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | | | |
Db 8 LRLGAPA 14

RESULT 5
US-09-171-705-38
; Sequence 38, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
; APPLICANT: BOOTS, ANNA M.H.
; APPLICANT: VERHEIJDEN, GILBERTUS F.M.
; TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
; TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
; FILE REFERENCE: O/96198 US
; CURRENT APPLICATION NUMBER: US/09/171,705
; CURRENT FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 72
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
; OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-72

Query Match 100.0%; Score 34; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 1.6e+05;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | | | |
Db 2 LRLGAPA 8

RESULT 4
US-09-171-705-37
; Sequence 37, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
; APPLICANT: BOOTS, ANNA M.H.
; APPLICANT: VERHEIJDEN, GILBERTUS F.M.
; TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
; TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
; FILE REFERENCE: O/96198 US
; CURRENT APPLICATION NUMBER: US/09/171,705
; CURRENT FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 37
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
; OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-37

Query Match 100.0%; Score 34; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | | | |
Db 8 LRLGAPA 14

RESULT 5
US-09-171-705-38
; Sequence 38, Application US/09171705
; Patent No. 6184204
; GENERAL INFORMATION:
; APPLICANT: BOOTS, ANNA M.H.
; APPLICANT: VERHEIJDEN, GILBERTUS F.M.
; TITLE OF INVENTION: NOVEL PEPTIDES SUITABLE FOR USE IN ANTIGEN SPECIFIC
; TITLE OF INVENTION: IMMUNOSUPPRESSIVE THERAPY
; FILE REFERENCE: O/96198 US
; CURRENT APPLICATION NUMBER: US/09/171,705
; CURRENT FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 38
LENGTH: 16
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:

OTHER INFORMATION: Description of Artificial Sequence: DERIVED FROM
OTHER INFORMATION: SEQUENCE OF HUMAN CARTILAGE (HC) -39 PROTEIN
US-09-171-705-38

Query Match 100.0%; Score 34; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 0.29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAP 7
Db 2 LRLGAP 8

RESULT 6

US-08-473-981A-8
Sequence 8, Application US/08473981A
Patent No. 5629162
GENERAL INFORMATION:
APPLICANT: defougercolles, Antonin R
APPLICANT: Springer, Timothy A
TITLE OF INVENTION: METHODS OF IDENTIFYING AGENTS WHICH MODULATE
TITLE OF INVENTION: ICAM-3 BINDING TO LFA-1 (AS AMENDED)
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W. SUITE 600
CITY: WASHINGTON
STATE: D. C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/473,981A
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: MILLONIG, ROBERT C
REGISTRATION NUMBER: 34,395
REFERENCE/DOCKET NUMBER: 1011.0560004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-473-981A-8

Query Match 88.2%; Score 30; DB 1; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAP 6
Db 13 LRLGAP 18

RESULT 7

US-08-474-087-8

Sequence 8, Application US/08474087
Patent No. 5891841
GENERAL INFORMATION:
APPLICANT: de Fougercolles, Antonin R
APPLICANT: Springer, Timothy A
TITLE OF INVENTION: METHODS OF USING INTERCELLULAR ADHESION MOLECULE-
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX, P.L.L.C.
STREET: 1100 NEW YORK AVENUE, N.W. SUITE 600
CITY: WASHINGTON
STATE: D. C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/474,087
FILING DATE: 07-JUN-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/038,990
FILING DATE: 23-DEC-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/712,879
FILING DATE: 11-JUN-1991
ATTORNEY/AGENT INFORMATION:
NAME: MILLONIG, ROBERT C
REGISTRATION NUMBER: 34,395
REFERENCE/DOCKET NUMBER: 1011.0560003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 56 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-474-087-8

Query Match 88.2%; Score 30; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAP 6
Db 13 LRLGAP 18

RESULT 8

US-08-466-265-3
Sequence 3, Application US/08466265
Patent No. 5712114
GENERAL INFORMATION:
APPLICANT: Mankovich, John A.
APPLICANT: Hammill, Linda
TITLE OF INVENTION: Compositions for Expression of Proteins in Host
TITLE OF INVENTION: Cells Using a Preprocollagen Signal Sequence
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: LAHIVE & COCKFIELD
STREET: 60 State Street, suite 510
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109-1875
COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,265
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 17
; OTHER INFORMATION: /note= "Xaa is Val or Ile"
; US-08-466-265-3

Query Match      82.4%; Score 28; DB 1; Length 23;
Best Local Similarity 83.3%; Pred. No. 8.1;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAP 6
   :|||||
DB 2 IRLGAP 7

RESULT 9
US-08-775-009-37
; Sequence 37, Application US/08775009
; Patent No. 5935783
; GENERAL INFORMATION:
; APPLICANT: Gong, Weilong
; APPLICANT: Emanuel, Beverly S.
; APPLICANT: Budarf, Marcia L.
; APPLICANT: Roe, Bruce
; TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
; TITLE OF INVENTION: Vellocardiofacial Syndrome Minimal Critical Region
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5935783ris, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,009
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yatto
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CH-0681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/466,265
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-030
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 23 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Modified-site
; LOCATION: 17
; OTHER INFORMATION: /note= "Xaa is Val or Ile"
; US-08-466-265-3

Query Match      82.4%; Score 28; DB 1; Length 23;
Best Local Similarity 83.3%; Pred. No. 8.1;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAP 6
   :|||||
DB 2 IRLGAP 7

RESULT 9
US-08-775-009-37
; Sequence 37, Application US/08775009
; Patent No. 5935783
; GENERAL INFORMATION:
; APPLICANT: Gong, Weilong
; APPLICANT: Emanuel, Beverly S.
; APPLICANT: Budarf, Marcia L.
; APPLICANT: Roe, Bruce
; TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digeorge and
; TITLE OF INVENTION: Vellocardiofacial Syndrome Minimal Critical Region
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
; ADDRESSEE: No. 5935783ris, LLP
; STREET: One Liberty Place - 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: U.S.A.
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/775,009
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Trujillo, Doreen Yatto
; REGISTRATION NUMBER: 35,719
; REFERENCE/DOCKET NUMBER: CH-0681
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-3100

; TELEFAX: (215) 568-3439
; INFORMATION FOR SEQ ID NO: 37:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 205 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-775-009-37

Query Match      82.4%; Score 28; DB 2; Length 205;
Best Local Similarity 85.7%; Pred. No. 82;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAP 7
   :|||||
DB 102 LSLGAP 108

RESULT 10
US-09-248-335-36
; Sequence 36, Application US/09248335
; Patent No. 6096504
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEF, DANIEL
; TITLE OF INVENTION: PLANT GLUTATHIONE-S-TRANSFERASE ENZYMES
; FILE REFERENCE: CL-1128-A
; CURRENT APPLICATION NUMBER: US/09/248,335
; CURRENT FILING DATE: 1999-02-10
; EARLIER APPLICATION NUMBER: 08/924,759
; EARLIER FILING DATE: 1997-September-05
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: Microsoft Word Version 7.0A
; SEQ ID NO 36
; LENGTH: 300
; TYPE: PRT
; ORGANISM: maize
; US-09-248-335-36

Query Match      82.4%; Score 28; DB 3; Length 300;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAP 6
   :|||||
DB 4 LRLGAP 9

RESULT 11
US-08-682-193A-2
; Sequence 2, Application US/08682193A
; Patent No. 5776740
; GENERAL INFORMATION:
; APPLICANT: HATAKEYAMA, Kazuhisa
; APPLICANT: GOTO, Makoto
; APPLICANT: TERASAWA, Masato
; APPLICANT: YUKAWA, Hideaki
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF
; TITLE OF INVENTION: L-TRYPTOPHANE
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 WATER STREET
; CITY: BOSTON
; STATE: MA
; COUNTRY: USA
; ZIP: 02019
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS

```

```
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/682,193a
; FILING DATE: 17-JUL-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 7-1811730
; FILING DATE: 18-JUL-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: DAVID, RESNICK S
; REGISTRATION NUMBER: 34,235
; REFERENCE/DOCKET NUMBER: 46643
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; TELEX: 200291 STRE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-682-193a-2

Query Match      82.4%; Score 28; DB 1; Length 434;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LRLGAPA 7
Db 372 LRLGTPA 378

RESULT 12
US-08-679-405-2
; Sequence 2, Application US/08679405
; Patent No. 5866393
; GENERAL INFORMATION:
; APPLICANT: Fuglsang, Claus
; APPLICANT: Halkier, Torben M.
; APPLICANT: Oxenboll, Karen M.
; APPLICANT: Berka, Randy M.
; APPLICANT: Cherry, Joel
; TITLE OF INVENTION: Haloperoxidases from Curvularia
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5866393o No. 5866393disk of No. 5866393th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/679,405
; FILING DATE: July 9, 1996
; FILING DATE: July 9, 1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,194
; FILING DATE: July 14, 1995
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 434 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; US-08-682-193a-2
```

```
; FILING DATE: February 21, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4441.210.US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 867-0123
; TELEFAX: (212) 878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-679-405-2

Query Match      82.4%; Score 28; DB 2; Length 600;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 LRLGAPA 7
Db 376 LTLGAPA 382

RESULT 13
US-08-842-799-2
; Sequence 2, Application US/08842799
; Patent No. 5965418
; GENERAL INFORMATION:
; APPLICANT: Fuglsang, Claus
; APPLICANT: Halkier, Torben M.
; APPLICANT: Oxenboll, Karen M.
; APPLICANT: Berka, Randy M.
; APPLICANT: Cherry, Joel
; TITLE OF INVENTION: Haloperoxidases from Curvularia
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 5965418o No. 5965418disk of No. 5965418th America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/842,799
; FILING DATE: 16-APR-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/679,405
; FILING DATE: July 9, 1996
; APPLICATION NUMBER: 60/001,194
; FILING DATE: July 14, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/603,534
; FILING DATE: February 21, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4441.210.US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 867-0123
; TELEFAX: (212) 878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 amino acids
```

```
;
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-842-799-2

Query Match      82.4%; Score 28; DB 2; Length 600;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
   | | | | |
Db 376 LTLGAPA 382

RESULT 14
US-09-271-778-2
; Sequence 2, Application US/0921778
; Patent No. 6221821
; GENERAL INFORMATION:
; APPLICANT: Svendsen, Allan
; APPLICANT: Jorgensen, Louise
; TITLE OF INVENTION: Haloperoxidases With Altered pH Profiles
; FILE REFERENCE: 5516.200-US
; CURRENT APPLICATION NUMBER: US/09/271,778
; CURRENT FILING DATE: 1999-03-18
; EARLIER APPLICATION NUMBER: PA 1998 00374
; EARLIER FILING DATE: 1998-03-18
; EARLIER APPLICATION NUMBER: 60/079,228
; EARLIER FILING DATE: 1998-03-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 600
; TYPE: PRT
; ORGANISM: Curvularia sp.
US-09-271-778-2

Query Match      82.4%; Score 28; DB 5; Length 600;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
   | | | | |
Db 376 LTLGAPA 382

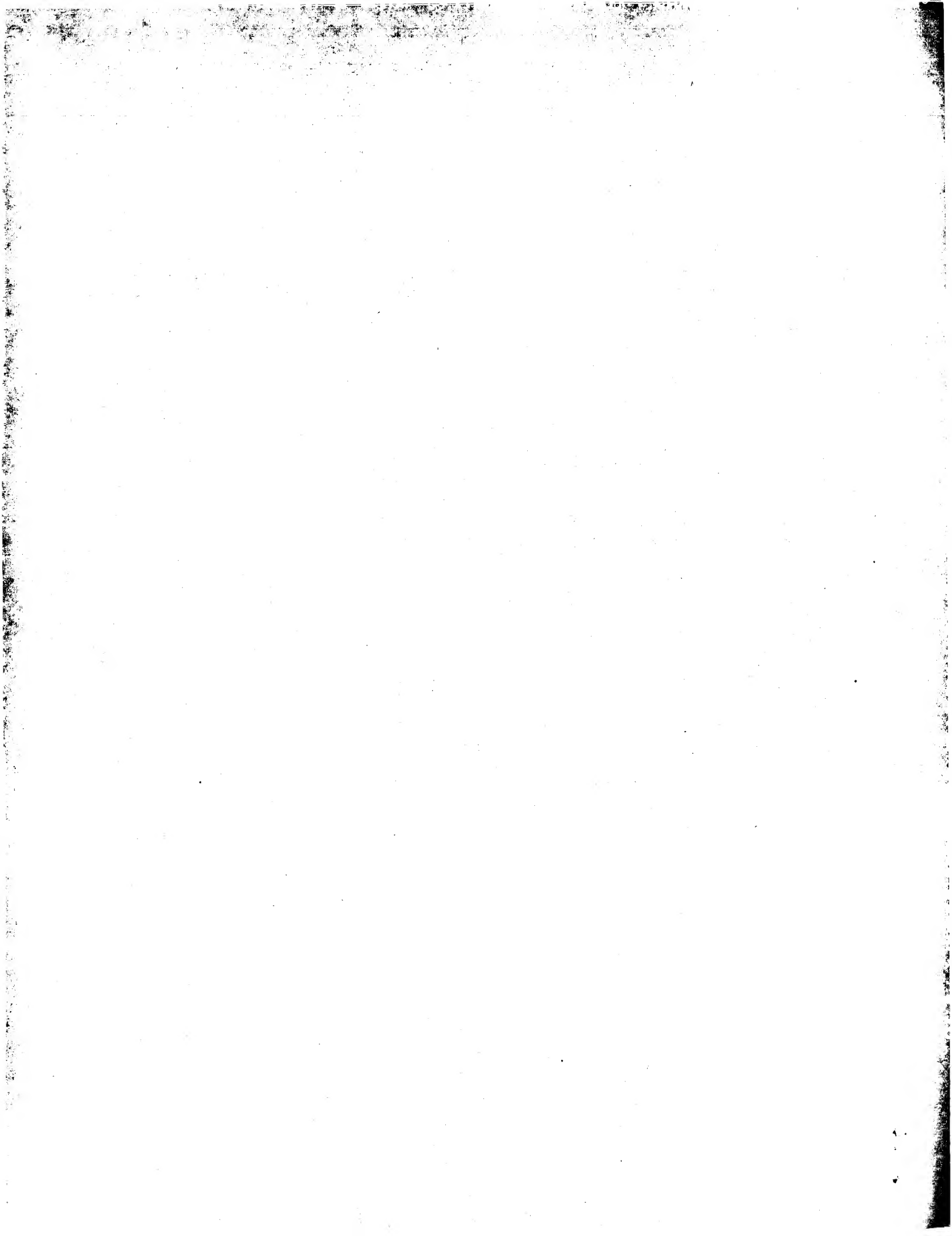
PCT-US96-11458-2
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/001,194
; FILING DATE: 14-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/603,534
; FILING DATE: 21-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Lambiris, Elias J.
; REGISTRATION NUMBER: 33,728
; REFERENCE/DOCKET NUMBER: 4441.204-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 867-0123
; TELEFAX: (212) 878-9655
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 600 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-11458-2

Search completed: December 7, 2001, 00:24:40
Job time: 2870 sec
```

```
Query Match      82.4%; Score 28; DB 4; Length 600;
Best Local Similarity 85.7%; Pred. No. 2.6e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
   | | | | |
Db 376 LTLGAPA 382

RESULT 15
PCT-US96-11458-2
; Sequence 2, Application PC/TUS9611458
; GENERAL INFORMATION:
; APPLICANT:
; APPLICANT:
; TITLE OF INVENTION: Haloperoxidases from Curvularia
; TITLE OF INVENTION: Verruculosa and Nucleic Acids Encoding Same
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Novo Nordisk of North America, Inc.
; STREET: 405 Lexington Avenue, Suite 6400
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10174
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/11458
; FILING DATE: 9-JUL-1996
; CLASSIFICATION:
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:32:57 ; Search time 401.11 Seconds
(without alignments)
4.846 Million cell updates/sec

Title: US-09-164-862B-3
Perfect score: 34
Sequence: 1 LRLGAPA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 3148936 seqs, 277657034 residues

Total number of hits satisfying chosen parameters: 3148936

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

- 1: /cgn2_6/ptodata/1/paa/US06_COMB.pap.*
- 2: /cgn2_6/ptodata/1/paa/US06_COMB.pap.*
- 3: /cgn2_6/ptodata/1/paa/US07_COMB.pap.*
- 4: /cgn2_6/ptodata/1/paa/US08_COMB.pap.*
- 5: /cgn2_6/ptodata/1/paa/US081_COMB.pap.*
- 6: /cgn2_6/ptodata/1/paa/US082_COMB.pap.*
- 7: /cgn2_6/ptodata/1/paa/US083_COMB.pap.*
- 8: /cgn2_6/ptodata/1/paa/US084_COMB.pap.*
- 9: /cgn2_6/ptodata/1/paa/US085_COMB.pap.*
- 10: /cgn2_6/ptodata/1/paa/US086_COMB.pap.*
- 11: /cgn2_6/ptodata/1/paa/US087_COMB.pap.*
- 12: /cgn2_6/ptodata/1/paa/US088_COMB.pap.*
- 13: /cgn2_6/ptodata/1/paa/US089_COMB.pap.*
- 14: /cgn2_6/ptodata/1/paa/US090_COMB.pap.*
- 15: /cgn2_6/ptodata/1/paa/US091_COMB.pap.*
- 16: /cgn2_6/ptodata/1/paa/US092_COMB.pap.*
- 17: /cgn2_6/ptodata/1/paa/US093_COMB.pap.*
- 18: /cgn2_6/ptodata/1/paa/US094_COMB.pap.*
- 19: /cgn2_6/ptodata/1/paa/US095_COMB.pap.*
- 20: /cgn2_6/ptodata/1/paa/US096_COMB.pap.*
- 21: /cgn2_6/ptodata/1/paa/US097_COMB.pap.*
- 22: /cgn2_6/ptodata/1/paa/US098_COMB.pap.*
- 23: /cgn2_6/ptodata/1/paa/US099_COMB.pap.*
- 24: /cgn2_6/ptodata/1/paa/US60_COMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------------------|--------------------|
| 1 | 34 | 100.0 | 7 | US-08-089-989-3 | Sequence 3, Appl1 |
| 2 | 34 | 100.0 | 7 | US-09-164-862B-3 | Sequence 3, Appl1 |
| 3 | 34 | 100.0 | 7 | US-09-215-077A-3 | Sequence 3, Appl1 |
| 4 | 34 | 100.0 | 114 | US-08-390-207-54 | Sequence 54, Appl1 |
| 5 | 34 | 100.0 | 151 | US-09-617-681A-1343 | Sequence 1343, Ap |
| 6 | 34 | 100.0 | 347 | US-09-760-475-2628 | Sequence 2628, Ap |
| 7 | 34 | 100.0 | 362 | PCT-US01-14827-14282 | Sequence 14282, A |
| 8 | 34 | 100.0 | 376 | PCT-US01-01239-1231 | Sequence 1231, Ap |
| 9 | 34 | 100.0 | 376 | US-09-764-902-1231 | Sequence 1231, Ap |

| | | | | | | |
|----|----|-------|-----|----|----------------------|--------------------|
| 10 | 34 | 100.0 | 383 | 11 | US-08-755-051-3 | Sequence 3, Appl1 |
| 11 | 34 | 100.0 | 383 | 15 | US-09-171-562-2 | Sequence 2, Appl1 |
| 12 | 34 | 100.0 | 383 | 18 | US-09-459-749B-17 | Sequence 17, Appl1 |
| 13 | 34 | 100.0 | 404 | 1 | PCT-US01-14827-14283 | Sequence 14283, A |
| 14 | 32 | 94.1 | 123 | 16 | US-09-248-796-17519 | Sequence 17519, A |
| 15 | 32 | 94.1 | 233 | 1 | PCT-US01-20638-284 | Sequence 284, App |
| 16 | 32 | 94.1 | 233 | 22 | US-09-893-737-284 | Sequence 284, App |
| 17 | 32 | 94.1 | 496 | 24 | US-60-324-109-20272 | Sequence 20272, A |
| 18 | 31 | 91.2 | 65 | 22 | US-09-867-550-1288 | Sequence 1288, Ap |
| 19 | 31 | 91.2 | 115 | 1 | PCT-US00-00466-478 | Sequence 478, App |
| 20 | 31 | 91.2 | 115 | 1 | PCT-US00-00466A-478 | Sequence 478, App |
| 21 | 31 | 91.2 | 115 | 18 | US-09-479-221-478 | Sequence 478, App |
| 22 | 31 | 91.2 | 115 | 19 | US-09-573-655A-825 | Sequence 825, App |
| 23 | 31 | 91.2 | 117 | 1 | PCT-US00-00466-477 | Sequence 477, App |
| 24 | 31 | 91.2 | 117 | 1 | PCT-US00-00466A-477 | Sequence 477, App |
| 25 | 31 | 91.2 | 117 | 18 | US-09-479-221-477 | Sequence 477, App |
| 26 | 31 | 91.2 | 133 | 21 | US-09-739-449-12718 | Sequence 12718, A |
| 27 | 31 | 91.2 | 133 | 22 | US-09-803-110-12718 | Sequence 12718, A |
| 28 | 31 | 91.2 | 326 | 17 | US-09-369-955-9 | Sequence 9, Appl1 |
| 29 | 31 | 91.2 | 387 | 1 | PCT-US01-08631-36343 | Sequence 36343, A |
| 30 | 30 | 88.2 | 56 | 21 | US-09-758-466-515 | Sequence 515, App |
| 31 | 30 | 88.2 | 123 | 20 | US-09-617-681A-1743 | Sequence 1743, Ap |
| 32 | 30 | 88.2 | 139 | 1 | PCT-US01-14827-8155 | Sequence 8155, Ap |
| 33 | 30 | 88.2 | 164 | 18 | US-09-417-507-30971 | Sequence 30971, A |
| 34 | 30 | 88.2 | 176 | 1 | PCT-US01-14827-14785 | Sequence 14785, A |
| 35 | 30 | 88.2 | 223 | 16 | US-09-252-991A-22261 | Sequence 22261, A |
| 36 | 30 | 88.2 | 232 | 23 | US-09-902-540-16617 | Sequence 16617, A |
| 37 | 30 | 88.2 | 238 | 24 | US-60-167-217-14100 | Sequence 14100, A |
| 38 | 30 | 88.2 | 238 | 24 | US-60-191-637-14066 | Sequence 14066, A |
| 39 | 30 | 88.2 | 245 | 18 | US-09-417-507-40422 | Sequence 40422, A |
| 40 | 30 | 88.2 | 263 | 18 | US-09-489-039A-14333 | Sequence 14333, A |
| 41 | 30 | 88.2 | 300 | 15 | US-09-107-532-5624 | Sequence 5624, Ap |
| 42 | 30 | 88.2 | 301 | 15 | US-09-107-532A-5624 | Sequence 5624, Ap |
| 43 | 30 | 88.2 | 301 | 1 | PCT-US00-05989-546 | Sequence 546, App |
| 44 | 30 | 88.2 | 301 | 23 | US-09-925-297-546 | Sequence 546, App |
| 45 | 30 | 88.2 | 307 | 16 | US-09-252-991A-22508 | Sequence 22508, A |

ALIGNMENTS

RESULT 1
US-08-089-989-3
; Sequence 3, Application US/08089989
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/089,989
; FILING DATE: 09-JUL-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: PD2759
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110

us-09-164-862b-3.rapm

Tue Dec 11 08:46:48 2001

```
;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: YKL-40 INTERNAL PEPTIDE B
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..7
; US-08-089-989-3

Query Match 100.0%; Score 34; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 1 LRLGAPA 7

RESULT 4
US-08-390-207-54
; Sequence 54, Application US/08390207
; GENERAL INFORMATION:
; APPLICANT: KATO, Suwan
; APPLICANT: OH, Suwan
; APPLICANT: SEKINE, Shingo
; APPLICANT: KIM, Namsoon
; APPLICANT: KATO, Takae
; APPLICANT: IWAHORI, Akiyo
; TITLE OF INVENTION: HUMAN CDNAS AND PROTEINS ENCODED THEREBY
; NUMBER OF SEQUENCES: 253
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Foley & Lardner
; STREET: 3000 K Street, N.W., Suite 500
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,207
; FILING DATE: 16-FEB-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-208077
; FILING DATE: 04-AUG-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 4-327619
; FILING DATE: 13-NOV-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 5-61431
; FILING DATE: 26-FEB-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/JP93/01095
; FILING DATE: 04-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/379,441
; FILING DATE: 03-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: WEGNER, Harold C.
; REGISTRATION NUMBER: 25,258
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202)672-5300
; TELEFAX: (202)672-5399
; TELEX: 904136
; INFORMATION FOR SEQ ID NO: 54:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 114 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-390-207-54

;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; IMMEDIATE SOURCE:
; CLONE: YKL-40 INTERNAL PEPTIDE B
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..7
; US-08-089-989-3

Query Match 100.0%; Score 34; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 1 LRLGAPA 7

RESULT 2
US-09-164-862B-3
; Sequence 3, Application US/09164862B
; GENERAL INFORMATION:
; APPLICANT: Price, Paul
; APPLICANT: JOHANSEN, Julia
; TITLE OF INVENTION: YKL-40 AS A MARKER AND PROGNOSTIC INDICATOR FOR CANCERS
; FILE REFERENCE: 407T-8955-00US
; CURRENT APPLICATION NUMBER: US/09/164,862B
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-164-862B-3

Query Match 100.0%; Score 34; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 2.9e+06;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 1 LRLGAPA 7

RESULT 3
US-09-215-077A-3
; Sequence 3, Application US/09215077A
; GENERAL INFORMATION:
; APPLICANT: PRICE, PAUL A.
; APPLICANT: JOHANSEN, JULIA S.
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR DEGRADATION OF
; FILE REFERENCE: 407T-895411US
; CURRENT APPLICATION NUMBER: US/09/215,077A
; CURRENT FILING DATE: 1998-12-18
; PRIOR APPLICATION NUMBER: 08/581,527
; PRIOR FILING DATE: 1996-04-17
; PRIOR APPLICATION NUMBER: 08/089,989
; PRIOR FILING DATE: 1993-07-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
```

Query Match 100.0%; Score 34; DB 7; Length 114;
Best Local Similarity 100.0%; Pred. No. 75;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
Db 34 LRLGAPA 40

RESULT 5
US-09-617-681A-1343
; Sequence 1343, Application US/09617681A
; GENERAL INFORMATION:
; APPLICANT: N. ALEXANDROV et al.
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1064P
; CURRENT APPLICATION NUMBER: US/09/617,681A
; CURRENT FILING DATE: 2000-07-19
; NUMBER OF SEQ ID NOS: 9311
; SEQ ID NO 1343
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..151
; OTHER INFORMATION: Ceres Seq. ID 1427951
; NAME/KEY: misc.feature
; LOCATION: 1..151
; OTHER INFORMATION: Xaa is any amino acid
US-09-617-681A-1343

Query Match 100.0%; Score 34; DB 20; Length 151;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
Db 36 LRLGAPA 42

RESULT 6
US-09-760-475-2628
; Sequence 2828, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2628
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (10)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (109)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-475-2628

Query Match 100.0%; Score 34; DB 21; Length 347;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
Db 209 LRLGAPA 215

RESULT 7
PCT-US01-14827-14282
; Sequence 14282, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 14282
; LENGTH: 362
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (1)..(336)
; OTHER INFORMATION: Glycosyl hydrolases family 18 domain identified by Pfam,
; OTHER INFORMATION: accession name Glyco_hydro_18, E-value=9.3e-184, Pfam score of
PCT-US01-14827-14282

Query Match 100.0%; Score 34; DB 1; Length 362;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
Db 224 LRLGAPA 230

RESULT 8
PCT-US01-01239-1231
; Sequence 1231, Application PC/TUS0101239
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT213PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01239
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1231
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-01239-1231

Query Match 100.0%; Score 34; DB 1; Length 376;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
Db 296 LRLGAPA 302

RESULT 9
US-09-764-902-1231
; Sequence 1231, Application US/09764902
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.

Tue Dec 11 08:46:48 2001

us-09-164-862b-3.rapm

;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
;; FILE REFERENCE: PT213
;; CURRENT APPLICATION NUMBER: US/09/764,902
;; CURRENT FILING DATE: 2001-01-17
;; Prior application data removed - refer to PALM or file wrapper
;; NUMBER OF SEQ ID NOS: 2318
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 1231
;; LENGTH: 376
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-09-764-902-1231

Query Match 100.0%; Score 34; DB 21; Length 376;
Best Local Similarity 100.0%; Pred. No. 2.6e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0;

QY 1 LRLGAPA 7
| | | | |

Db 296 LRLGAPA 302

RESULT 10
US-08-755-051-3
; Sequence 3, Application US/08755051
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Au-Young, Janice
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN CHITOSIDASE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: US
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/755,051
; FILING DATE: Herewith
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0152 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 383 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 398912
US-08-755-051-3

Query Match 100.0%; Score 34; DB 11; Length 383;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 LRLGAPA 7
| | | | |
Db 245 LRLGAPA 251

RESULT 11
US-09-171-562-2
; Sequence 2, Application US/09171562
; GENERAL INFORMATION:
; APPLICANT: Boots, Anna M.H.
; APPLICANT: Verheijden, Gilbertus F.M.
; APPLICANT: Bos, Ebo S.
; TITLE OF INVENTION: Autoantigen and Proteins Structurally Related thereto
; TITLE OF INVENTION: for Use in Immunotherapy of Autoimmune Diseases
; FILE REFERENCE: O/961920S
; CURRENT APPLICATION NUMBER: US/09/171,562
; CURRENT FILING DATE: 1998-10-19
; EARLIER APPLICATION NUMBER: PCT/EP97/01903
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-171-562-2

Query Match 100.0%; Score 34; DB 15; Length 383;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
Db 245 LRLGAPA 251

RESULT 12
US-09-459-749B-17
; Sequence 17, Application US/09459749B
; GENERAL INFORMATION:
; APPLICANT: Millis, Albert J. T.
; TITLE OF INVENTION: Compositions and Methods For Altering Cell Migration
; FILE REFERENCE: 0794,016A
; CURRENT APPLICATION NUMBER: US/09/459,749B
; CURRENT FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: 60/111,856
; PRIOR FILING DATE: 1998-12-11
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 17
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-459-749B-17

Query Match 100.0%; Score 34; DB 18; Length 383;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
Db 245 LRLGAPA 251

RESULT 13
PCT-US01-14827-14283
; Sequence 14283, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hysseq, Inc

```
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 14283
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: DOMAIN
; LOCATION: (153)..(165)
; OTHER INFORMATION: Chitinases family 18 proteins domain identified by EMATRIX,
; OTHER INFORMATION: accession number BL01095B, p-value=2.469e-09, raw score of 10.82
; NAME/KEY: DOMAIN
; LOCATION: (43)..(379)
; OTHER INFORMATION: Glycosyl hydrolases family 18 domain identified by Pfam,
; OTHER INFORMATION: accession name Glyco_hydro_18, E-value=3.1e-187, Pfam score of
; OTHER INFORMATION: 630.9
; PCT-US01-14827-14283
```

```
Query Match 100.0%; Score 34; DB 1; Length 404;
Best Local Similarity 100.0%; Pred. No. 2.8e+02; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LRLGAPA 7
Db 266 LRLGAPA 272
|||||
```

```
RESULT 14
US-09-248-796-17519
; Sequence 17519, Application US/09248796
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796
; CURRENT FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 28206
; SEQ ID NO 17519
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796-17519
```

```
Query Match 94.1%; Score 32; DB 16; Length 123;
Best Local Similarity 85.7%; Pred. No. 2.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LRLGAPA 7
Db 28 LRLGAPA 34
|||||
```

```
RESULT 15
PCT-US01-20638-284
; Sequence 284, Application PC/TUS0120638
; GENERAL INFORMATION:
; APPLICANT: ZymoGenetics, Inc.
; TITLE OF INVENTION: MAMMALIAN SECRETED PROTEINS
; FILE REFERENCE: 00-41PC
; CURRENT APPLICATION NUMBER: PCT/US01/20638
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: US 60/215,446
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 329
```

```
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 284
; LENGTH: 233
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-20638-284
```

```
Query Match 94.1%; Score 32; DB 1; Length 233;
Best Local Similarity 85.7%; Pred. No. 4.1e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 LRLGAPA 7
Db 86 LRLGAPA 92
|||||
```

```
Search completed: December 7, 2001, 00:32:58
Job time: 797 sec
```

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:35:57 ; Search time 169.18 Seconds
(without alignments)
0.419 Million cell updates/sec

Title: US-09-164-862b-3
Perfect score: 34
Sequence: 1 LRLGAPA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 56692 seqs, 10114640 residues

Total number of hits satisfying chosen parameters: 56692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_New.*
1: /cgn2_6/ptodata/2/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/2/paa/US05_NEW_COMB.pep.*
3: /cgn2_6/ptodata/2/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/2/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/2/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/2/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 34 | 100.0 | 10 | 5 | US-09-572-404B-2228 |
| 2 | 29 | 85.3 | 347 | 5 | US-09-620-394B-4891 |
| 3 | 29 | 85.3 | 471 | 5 | US-09-620-394B-4890 |
| 4 | 29 | 85.3 | 497 | 5 | US-09-620-394B-4889 |
| 5 | 28 | 82.4 | 400 | 5 | US-09-620-394B-6325 |
| 6 | 28 | 82.4 | 412 | 5 | US-09-815-242-5451 |
| 7 | 28 | 82.4 | 412 | 5 | US-09-815-242-12381 |
| 8 | 28 | 82.4 | 412 | 5 | US-09-815-242-12034 |
| 9 | 28 | 82.4 | 418 | 5 | US-09-620-394B-6324 |
| 10 | 28 | 82.4 | 439 | 5 | US-09-620-394B-6323 |
| 11 | 28 | 82.4 | 486 | 5 | US-09-815-242-13285 |
| 12 | 28 | 82.4 | 614 | 5 | US-09-832-614A-2 |
| 13 | 28 | 82.4 | 4342 | 5 | US-09-815-242-5107 |
| 14 | 27 | 79.4 | 356 | 5 | US-09-897-516-6508 |
| 15 | 27 | 79.4 | 419 | 5 | US-09-815-242-13792 |
| 16 | 27 | 79.4 | 502 | 5 | US-09-620-394B-2083 |
| 17 | 27 | 79.4 | 520 | 5 | US-09-620-394B-2082 |
| 18 | 27 | 79.4 | 529 | 5 | US-09-620-394B-2081 |
| 19 | 27 | 79.4 | 4999 | 5 | US-09-976-059-14 |
| 20 | 27 | 79.4 | 4999 | 5 | US-09-976-059-15 |
| 21 | 26 | 76.5 | 130 | 5 | US-09-826-734-162 |
| 22 | 26 | 76.5 | 134 | 5 | US-09-620-394B-3902 |
| 23 | 26 | 76.5 | 162 | 5 | US-09-826-734-112 |
| 24 | 26 | 76.5 | 162 | 5 | US-09-826-734-118 |
| 25 | 26 | 76.5 | 189 | 5 | US-09-620-394B-3901 |
| 26 | 26 | 76.5 | 208 | 5 | US-09-620-394B-3900 |
| 27 | 26 | 76.5 | 248 | 5 | US-09-897-516-7983 |

28 26 76.5 331 5 US-09-897-516-7049 Sequence 7049, Ap
29 26 76.5 348 5 US-09-620-394B-452 Sequence 452, App
30 26 76.5 357 5 US-09-620-394B-451 Sequence 451, App
31 26 76.5 363 5 US-09-620-394B-450 Sequence 450, App
32 26 76.5 368 5 US-09-170-496D-20 Sequence 20, App
33 26 76.5 368 5 US-09-170-496D-174 Sequence 174, App
34 26 76.5 412 5 US-09-815-242-10827 Sequence 10827, A
35 26 76.5 427 5 US-09-815-242-10858 Sequence 10858, A
36 26 76.5 459 5 US-09-129-112-9 Sequence 9, Appl
37 26 76.5 653 5 US-09-639-696B-5 Sequence 5, Appl
38 26 76.5 653 5 US-09-639-696C-5 Sequence 5, Appl
39 26 76.5 828 1 PCT-US01-27760-706 Sequence 706, App
40 26 76.5 1005 5 US-09-988-117-1 Sequence 1, Appl
41 26 76.5 1045 5 US-09-555-342A-2 Sequence 2, Appl
42 25 73.5 30 5 US-09-922-261-34 Sequence 34, Appl
43 25 73.5 113 5 US-09-620-394B-5449 Sequence 5449, Ap
44 25 73.5 123 5 US-09-770-834-10 Sequence 10, Appl
45 25 73.5 150 5 US-09-828-792-730 Sequence 730, App

ALIGNMENTS

RESULT 1
US-09-572-404B-2228
; Sequence 2228, Application US/09572404B
; GENERAL INFORMATION:
; APPLICANT: Proteom Ltd
; TITLE OF INVENTION: Complementary peptide ligands from the human genome
; FILE REFERENCE: Human patent
; CURRENT APPLICATION NUMBER: US/09/572,404B
; CURRENT FILING DATE: 2000-05-17
; NUMBER OF SEQ ID NOS: 4203
; SOFTWARE: Protpatent version 1.0
; SEQ ID NO 2228
; LENGTH: 10
; TYPE: PRT
; ORGANISM: Homo Sapiens
; FEATURE:
; OTHER INFORMATION: sequence located in SHMT2 at 421-430 and may interact with Seq
; OTHER INFORMATION: in this patent.
US-09-572-404B-2228

Query Match 100.0%; Score 34; DB 5; Length 10;
Best Local Similarity 100.0%; Pred. No. 0.091;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 4 LRLGAPA 10
|||||||

RESULT 2
US-09-620-394B-4891
; Sequence 4891, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
; TITLE OF INVENTION: Thereby
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4891
; LENGTH: 347
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..347
; OTHER INFORMATION: Xaa is any amino acid

```

; NAME/KEY: misc_feature
; LOCATION: 1..347
; OTHER INFORMATION: Ceres Seq. ID 1393595
US-09-620-394B-4891

Query Match      85.3%; Score 29; DB 5; Length 347;
Best Local Similarity 71.4%; Pred. No. 31;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 264 VRIGAPA 270

RESULT 3
US-09-620-394B-4890
; Sequence 4890, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4890
; LENGTH: 471
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: misc_feature
; LOCATION: 1..471
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..471
; OTHER INFORMATION: Ceres Seq. ID 1393594
US-09-620-394B-4890

Query Match      85.3%; Score 29; DB 5; Length 471;
Best Local Similarity 71.4%; Pred. No. 42;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 388 VRIGAPA 394

RESULT 4
US-09-620-394B-4889
; Sequence 4889, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 4889
; LENGTH: 497
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: misc_feature
; LOCATION: 1..497
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..497
; OTHER INFORMATION: Ceres Seq. ID 1393593

US-09-620-394B-4889

Query Match      85.3%; Score 29; DB 5; Length 497;
Best Local Similarity 71.4%; Pred. No. 44;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 414 VRIGAPA 420

RESULT 5
US-09-620-394B-6325
; Sequence 6325, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 6325
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; NAME/KEY: misc_feature
; LOCATION: 1..400
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..400
; OTHER INFORMATION: Ceres Seq. ID 1395987
US-09-620-394B-6325

Query Match      82.4%; Score 28; DB 5; Length 400;
Best Local Similarity 85.7%; Pred. No. 58;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 239 LSLGAPA 245

RESULT 6
US-09-815-242-5451
; Sequence 5451, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Karl L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
```



```

; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5451
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-5451

```

```

Query Match      82.4%; Score 28; DB 5; Length 412;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LRLGAPA 7
   :||| ||
Db 357 IRLGTPA 363

```

```

RESULT 7
US-09-815-242-12381
; Sequence 12381, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Travick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12381
; LENGTH: 412
; TYPE: PRT
; ORGANISM: Staphylococcus aureus
US-09-815-242-12381

```

```

Query Match      82.4%; Score 28; DB 5; Length 412;
Best Local Similarity 71.4%; Pred. No. 60;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LRLGAPA 7
   :||| ||
Db 357 IRLGTPA 363

```

RESULT 8

```

US-09-815-242-12034
; Sequence 12034, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Travick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12034
; LENGTH: 417
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-12034

```

```

Query Match      82.4%; Score 28; DB 5; Length 417;
Best Local Similarity 71.4%; Pred. No. 61;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 LRLGAPA 7
   :||| ||
Db 361 LRLGTPA 367

```

```

RESULT 9
US-09-620-394B-6324
; Sequence 6324, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nickolai
; APPLICANT: BROVER, Vyacheslav
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptid
; FILE REFERENCE: Thereby
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 5131
; SEQ ID NO 6324
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1..418
; OTHER INFORMATION: xaa is any amino acid
; NAME/KEY: misc.feature
; LOCATION: 1..418
; OTHER INFORMATION: Ceres Seq. ID 1395986
US-09-620-394B-6324

```

Query Match 82.4%; Score 28; DB 5; Length 418;
Best Local Similarity 85.7%; Pred. No. 61;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
DB 257 LSLGAPA 263

RESULT 10
US-09-620-394B-6323
; Sequence 6323, Application US/09620394B
; GENERAL INFORMATION:
; APPLICANT: ALEXANDROV, Nikolai
; TITLE OF INVENTION: Sequence-Determined DNA Fragments and Corresponding Polypeptides
; FILE REFERENCE: 2750-1067P
; CURRENT APPLICATION NUMBER: US/09/620,394B
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 9131
; SEQ ID NO 6323
; LENGTH: 439
; TYPE: PRT
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..439
; OTHER INFORMATION: Xaa is any amino acid
; NAME/KEY: misc_feature
; LOCATION: 1..439
; OTHER INFORMATION: Ceres Seq. ID 1395985
US-09-620-394B-6323

Query Match 82.4%; Score 28; DB 5; Length 439;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 1; Indels 1; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
DB 278 LSLGAPA 284

RESULT 11
US-09-815-242-13285
; Sequence 13285, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13285
; LENGTH: 486
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-815-242-13285

Query Match 82.4%; Score 28; DB 5; Length 486;
Best Local Similarity 71.4%; Pred. No. 70;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
DB 234 LRLGAPA 240

RESULT 12
US-09-832-614A-2
; Sequence 2, Application US/09832614A
; GENERAL INFORMATION:
; APPLICANT: Novozymes A/S
; APPLICANT: Danielsen, Steffen
; APPLICANT: Schneider, Palle
; TITLE OF INVENTION: Nucleic acids encoding polypeptides having haloperoxidase acti
; FILE REFERENCE: 10173.200-US
; CURRENT APPLICATION NUMBER: US/09/832,614A
; CURRENT FILING DATE: 2001-04-11
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Geniculosporium sp.
US-09-832-614A-2

Query Match 82.4%; Score 28; DB 5; Length 614;
Best Local Similarity 85.7%; Pred. No. 88;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
| | | | |
DB 384 LTLGAPA 390

RESULT 13
US-09-815-242-5107
; Sequence 5107, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard
; TITLE OF INVENTION: Identification of Essential Genes in
; FILE REFERENCE: ELITRA.011a
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26

```

; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5107
; LENGTH: 4342
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-815-242-5107

```

```

Query Match      82.4%; Score 28; DB 5; Length 4342;
Best Local Similarity 83.3%; Pred. No. 5.8e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 LRLGAP 6
   ||:||||
Db 1996 LRLGAP 2001

```

```

RESULT 14
US-09-897-516-6508
; Sequence 6508, Application US/09897516
; GENERAL INFORMATION:
; APPLICANT: Corbin, David R.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Huesling, Joseph E.
; APPLICANT: Krasomil-Osterfeld, Karina C.
; APPLICANT: Malvar, Thomas M.
; APPLICANT: Slater, Steven C.
; APPLICANT: Spiridonov, Sergei
; TITLE OF INVENTION: Xenorhabdus sp. Genome Sequences And Uses Thereof
; FILE REFERENCE: 38-21(51847)B
; CURRENT APPLICATION NUMBER: US/09/897,516
; CURRENT FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: US 60/215, 161
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 8409
; SEQ ID NO 6508
; LENGTH: 356
; TYPE: PRT
; ORGANISM: Xenorhabdus sp.
US-09-897-516-6508

```

```

Query Match      79.4%; Score 27; DB 5; Length 356;
Best Local Similarity 71.4%; Pred. No. 84;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

Qy 1 LRLGAP 7
   | :||||
Db 322 LNLGAPA 328

```

```

RESULT 15
US-09-815-242-13792
; Sequence 13792, Application US/09815242
; GENERAL INFORMATION:
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari L.
; APPLICANT: Zyskind, Judith W.
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John D.
; APPLICANT: Carr, Grant J.
; APPLICANT: Yamamoto, Robert T.
; APPLICANT: Xu, H. Howard

```

```

; TITLE OF INVENTION: Identification of Essential Genes in
; FILE OF INVENTION: Prokaryotes
; FILE REFERENCE: ELITRA_011A
; CURRENT APPLICATION NUMBER: US/09/815,242
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; NUMBER OF SEQ ID NOS: 14110
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13792
; LENGTH: 419
; TYPE: PRT
; ORGANISM: Salmonella typhi
US-09-815-242-13792

```

```

Query Match      79.4%; Score 27; DB 5; Length 419;
Best Local Similarity 57.1%; Pred. No. 99;
Matches 4; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 1 LRLGAPA 7
   ||:||||
Db 364 IRLGSPA 370

```

```

Search completed: December 7, 2001, 00:35:58
Job time: 8/2 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:26:04 ; Search time 71.99 Seconds
(without alignments)
7.407 Million cell updates/sec

Title: US-09-164-862B-3

Perfect score: 34

Sequence: 1 LRLGAPA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR_68:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 34 | 100.0 | 383 | 2 S51327 | heparin-binding gl |
| 2 | 34 | 100.0 | 383 | 2 A49562 | cartilage glycopro |
| 3 | 34 | 100.0 | 474 | 2 B46746 | glycine hydroxymet |
| 4 | 34 | 100.0 | 475 | 2 A33696 | glycine hydroxymet |
| 5 | 31 | 91.2 | 117 | 2 T07645 | PEARL1 1 protein h |
| 6 | 31 | 91.2 | 427 | 2 S75210 | glycine hydroxymet |
| 7 | 31 | 91.2 | 664 | 2 S61231 | cyck protein - Rhi |
| 8 | 31 | 91.2 | 676 | 1 S54750 | cytochrome c-type |
| 9 | 30 | 88.2 | 237 | 2 S71022 | transcription init |
| 10 | 30 | 88.2 | 322 | 1 JQ1028 | 6-phosphofructokin |
| 11 | 30 | 88.2 | 414 | 1 JQ1016 | glycine hydroxymet |
| 12 | 30 | 88.2 | 414 | 2 H81383 | glycine hydroxymet |
| 13 | 30 | 88.2 | 423 | 2 C72561 | hypothetical prote |
| 14 | 30 | 88.2 | 424 | 2 E82743 | serine hydroxymeth |
| 15 | 30 | 88.2 | 432 | 1 S15203 | glycine hydroxymet |
| 16 | 30 | 88.2 | 463 | 2 A69905 | conserved hypothet |
| 17 | 30 | 88.2 | 472 | 2 T37918 | serine hydroxymeth |
| 18 | 30 | 88.2 | 481 | 2 T42226 | hypothetical prote |
| 19 | 30 | 88.2 | 483 | 1 A46746 | glycine hydroxymet |
| 20 | 30 | 88.2 | 484 | 1 XYRBS | glycine hydroxymet |
| 21 | 30 | 88.2 | 484 | 2 A40202 | glycine hydroxymet |
| 22 | 30 | 88.2 | 485 | 2 T38255 | hypothetical prote |
| 23 | 30 | 88.2 | 486 | 2 T35445 | probable integral |
| 24 | 30 | 88.2 | 514 | 2 T44976 | hydantoinase homol |
| 25 | 30 | 88.2 | 533 | 2 D83685 | hydantoinase BH188 |
| 26 | 30 | 88.2 | 663 | 2 D83545 | probable helicase |
| 27 | 30 | 88.2 | 756 | 2 T00367 | hypothetical prote |
| 28 | 30 | 88.2 | 758 | 2 H70649 | carbon starvation |
| 29 | 30 | 88.2 | 1016 | 2 T31343 | proline dehydrogen |

| | | | | | | |
|----|----|------|------|---|--------|--------------------|
| 30 | 30 | 88.2 | 1067 | 2 | T30061 | hypothetical prote |
| 31 | 30 | 88.2 | 1224 | 2 | S25952 | gene cob intron 3 |
| 32 | 30 | 88.2 | 1694 | 2 | A83512 | hypothetical prote |
| 33 | 30 | 88.2 | 2104 | 2 | H86127 | hypothetical prote |
| 34 | 29 | 85.3 | 111 | 2 | T05907 | glycine hydroxymet |
| 35 | 29 | 85.3 | 197 | 2 | B72623 | hypothetical prote |
| 36 | 29 | 85.3 | 206 | 2 | B75505 | conserved hypothet |
| 37 | 29 | 85.3 | 263 | 2 | G75590 | hypothetical prote |
| 38 | 29 | 85.3 | 331 | 2 | B83486 | cobalamin biosynth |
| 39 | 29 | 85.3 | 337 | 2 | C65028 | hypothetical 36.2K |
| 40 | 29 | 85.3 | 337 | 2 | F85895 | probable membrane |
| 41 | 29 | 85.3 | 381 | 2 | F75270 | cytochrome P450 - |
| 42 | 29 | 85.3 | 434 | 2 | S30334 | glycine hydroxymet |
| 43 | 29 | 85.3 | 469 | 2 | S61632 | glycine hydroxymet |
| 44 | 29 | 85.3 | 471 | 2 | B71400 | glycine hydroxymet |
| 45 | 29 | 85.3 | 922 | 2 | D75615 | excinuclease ABC c |

ALIGNMENTS

RESULT 1

S51327

heparin-binding glycoprotein 38K - pig

C:Species: Sus scrofa domestica (domestic pig)

C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 22-Jun-1999

C:Accession: S51327

R:Shackelton, L.M.; Mann, D.M.; Millis, A.J.T.

submitted to the EMBL Data Library, January 1995

A:Description: Identification of a 38kDa heparin-binding glycoprotein (gp38k) in diff

A:Reference number: S51327

A:Accession: S51327

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-383 <SHA>

A:Cross-references: EMBL:Z47803; NID:G634097; PIDN:CAA87764.1; PID:G634098

C:Superfamily: Streptomyces chitinase chi40

Query Match 100.0%; Score 34; DB 2; Length 383;

Best Local Similarity 100.0%; Pred. No. 7.9;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7

DB 245 LRLGAPA 251

RESULT 2

A49562

cartilage glycoprotein gp39 precursor - human

N:Alternate names: 39K synovial protein

C:Species: Homo sapiens (man)

C:Date: 23-Mar-1995 #sequence_revision 23-Mar-1995 #text_change 22-Jun-1999

C:Accession: A49562; S10677; A33162

R:Hakala, B.E.; White, C.; Recklies, A.D.

J. Biol. Chem. 268, 25803-25810, 1993

A:Title: Human cartilage gp-39, a major secretory product of articular chondrocytes a

A:Reference number: A49562; MUID:94084658

A:Accession: A49562

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-383 <HAK>

A:Cross-references: GB:M80927; NID:G348911; PIDN:AAA16074.1; PID:G348912

R:Nyirkos, P.; Golds, E.E.

Biochem. J. 269, 265-268, 1990

A:Title: Human synovial cells secrete a 39 kDa protein similar to a bovine mammary pr

A:Reference number: S10677; MUID:90328983

A:Accession: S10677

A:Molecule type: protein

A:Residues: 22-40,'X',42-45 <NY2>

C:Superfamily: Streptomyces chitinase chi40

C:Keywords: cartilage; extracellular protein; glycoprotein

submitted to the EMBL Data Library, July 1995

A:Description: Characterization of the cydHJKL genes involved in cytochrome c biogenesis
A:Reference number: S61229
A:Accession: S61231
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-664
A:Cross-references: EMBL:X89726; NID:g967064; PIDN:CAA61878.1; PID:g967067
C:Superfamily: nrfe protein

Query Match 91.2%; Score 31; DB 2; Length 664;
Best Local Similarity 85.7%; Pred. No. 62;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
||:||||
DB 644 LRVGAPA 650

RESULT 8

S54750
cytochrome c-type synthesis protein cycK - Rhizobium meliloti
C:Species: Rhizobium meliloti
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S54750; S49616
R:Kereszt, A.; Slaska-Kiss, K.; Putnok, P.; Banfalvi, Z.; Kondorosi, A.
Mol. Gen. Genet. 247, 39-47, 1995
A:Title: The cydHJKL genes of Rhizobium meliloti involved in cytochrome c biogenesis are
A:Reference number: S54748; MUID:95231514
A:Accession: S54750
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-676 <K&E>
A:Cross-references: EMBL:X82560; NID:g575369; PIDN:CAA57906.1; PID:g575372
C:Genetics:
A:Gene: cycK
C:Superfamily: nrfe protein

Query Match 91.2%; Score 31; DB 1; Length 676;
Best Local Similarity 85.7%; Pred. No. 64;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
||:||||
DB 642 LRVGAPA 648

RESULT 9

S71022
transcription initiation factor sigma, flagellar-specific - Yersinia enterocolitica
C:Species: Yersinia enterocolitica
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 08-Oct-1999
C:Accession: S71022; S73318
R:Kapatal, V.; Olson, J.W.; Pepe, J.C.; Miller, V.L.; Minnich, S.A.
Mol. Microbiol. 19, 1061-1071, 1996
A:Title: Temperature-dependent regulation of Yersinia enterocolitica class III flagellar
A:Reference number: S70934; MUID:96249698
A:Accession: S71022
A:Status: nucleic acid sequence not shown
A:Molecule type: DNA
A:Residues: 1-237 <K&E>
A:Cross-references: EMBL:L33466; NID:g904171; PIDN:AAB38522.1; PID:g904172
A:Experimental source: strain 8081; type 0:8
R:Kapatal, V.; Olson, J.W.; Pepe, J.C.; Miller, V.L.; Minnich, S.A.
submitted to the EMBL Data Library, August 1995
A:Reference number: S73318
A:Accession: S73318
A:Molecule type: DNA
A:Residues: 10-90, 'L', 92-183, 'I', 184-237 <KAW>
A:Cross-references: EMBL:L33466; NID:g904171; PIDN:AAB38522.1; PID:g904172
C:Genetics:

A:Gene: fliA

A:Start codon: GTG
C:Superfamily: transcription initiation factor sigma; transcription initiation factor
C:Keywords: DNA binding; sigma factor; transcription initiation

Query Match 88.2%; Score 30; DB 2; Length 237;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLGAAPA 7
|||||
DB 108 RLGAAPA 113

RESULT 10

JQ1028
6-phosphofructokinase (EC 2.7.1.11) 1 - Thermus aquaticus
N:Alternate names: phosphofructokinase 1; phosphohexokinase
C:Species: Thermus aquaticus
C:Date: 10-Mar-1994 #sequence_revision 14-Jul-1994 #text_change 11-Jun-1999
C:Accession: JQ1028
R:Xu, J.; Seki, M.; Denda, K.; Yoshida, M.
Biochem. Biophys. Res. Commun. 176, 1313-1318, 1991
A:Title: Molecular cloning of phosphofructokinase 1 gene from a thermophilic bacteriu
A:Reference number: JQ1028; MUID:91248220
A:Accession: JQ1028
A:Molecule type: DNA
A:Residues: 1-322 <XU>
A:Cross-references: GB:M71213; NID:g155116; PIDN:AAA27501.1; PID:g155117
C:Genetics:
A:Gene: pfk1
C:Superfamily: 6-phosphofructokinase; 6-phosphofructokinase 1 homology
C:Keywords: allosteric regulation; ATP; glycolysis; phosphotransferase
F:4-281/Domain: 6-phosphofructokinase 1 homology <6PFL>

Query Match 88.2%; Score 30; DB 1; Length 322;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLGAAPA 7
|||||
DB 269 RLGAAPA 274

RESULT 11

JQ1016
glycine hydroxymethyltransferase (EC 2.1.2.1) - Campylobacter jejuni
N:Alternate names: serine aldolase; serine hydroxymethylase; serine methylase; threon
C:Species: campylobacter jejuni
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 29-Sep-1999
C:Accession: JQ1016; S16468
R:Chan, V.L.; Bingham, H.L.
Gene 101, 51-58, 1991
A:Title: Complete sequence of the Campylobacter jejuni glyA gene encoding serine hydr
A:Reference number: JQ1016; MUID:91285434
A:Accession: JQ1016
A:Molecule type: DNA
A:Residues: 1-414 <CHAL>
A:Cross-references: EMBL:X53816; NID:g40533; PIDN:CAA37812.1; PID:g40534
R:Chan, V.L.; Bingham, H.L.
Gene 101, 51-58, 1990
A:Title: Complete sequence of the Campylobacter jejuni glyA gene encoding serine hydr
A:Reference number: S16468; MUID:91285434
A:Accession: S16468
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-414 <CHAZ>
A:Cross-references: EMBL:X53816; NID:g40533; PIDN:CAA37812.1; PID:g40534
C:Comment: In the presence of tetrahydrofolate, this enzyme catalyzes the reversible
C:Genetics:
A:Gene: glyA

C:Superfamily: glycine hydroxymethyltransferase
C:Keywords: phosphoprotein; pyridoxal phosphate; transferase
F:224/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 88.2%; Score 30; DB 1; Length 414;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 355 LRLGTPA 361
|||||
RESULT 12
H81383
glycine hydroxymethyltransferase (EC 2.1.2.1) Cj0402 [imported] - Campylobacter jejuni
C:Species: Campylobacter jejuni
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 08-Sep-2000
C:Accession: H81383
R:Parkhill, J.; Wren, B.W.; Mungall, K.; Ketley, J.M.; Churcher, C.; Basham, D.; Chillingworth, C.W.; Quail, M.; Rajandream, M.A.; Rutherford, K.M.; VanVliet, A.; Whitehead, S.; Barrer, Nature 403, 665-668, 2000
A:Title: The genome sequence of the food-borne pathogen Campylobacter jejuni reveals hypA
A:Reference number: A81250; MUID:20150912
A:Accession: H81383
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-414 <PAR>
A:Cross-references: GB:AL139075; GB:AL111168; NID:g6967817; PIDN:CAB74238.1; PID:g696787
A:Experimental source: serotype O2, strain NCTC 11168
C:Genetics:
A:Gene: glyA; Cj0402
C:Superfamily: glycine hydroxymethyltransferase
C:Keywords: transferase

Query Match 88.28; Score 30; DB 2; Length 414;
Best Local Similarity 85.7%; Pred. No. 66;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 355 LRLGTPA 361
|||||
RESULT 13
C72561
hypothetical protein APE1773 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 20-Jun-2000
C:Accession: C72561
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takai, A.; H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A:Reference number: A72450; MUID:99310339
A:Accession: C72561
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-423 <KAW>
A:Cross-references: DBJ:AP000062; NID:g5105244; PIDN:BA080776.1; PID:g5105463
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1773
C:Superfamily: Methanococcus jannaschii conserved hypothetical protein MJ0709

Query Match 88.2%; Score 30; DB 2; Length 423;
Best Local Similarity 100.0%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLGAPA 7

Db 162 RLGAPA 167
|||||

RESULT 14
E82743
serine hydroxymethyltransferase XF0946 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Sep-2000
C:Accession: E82743
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.
A:Reference number: A82515; MUID:20365717
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82743
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-424 <SIM>
A:Cross-references: GB:AE003933; GB:AE003849; NID:g9105863; PIDN:AAF83756.1; GSPDB:GN
A:Experimental source: strain 9a5c
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R. Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier as-Neco, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to Genbank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.D.; Marques, M.V.; Martins A:Authors: Martins, A.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C. F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L. A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF0946
C:Superfamily: glycine hydroxymethyltransferase

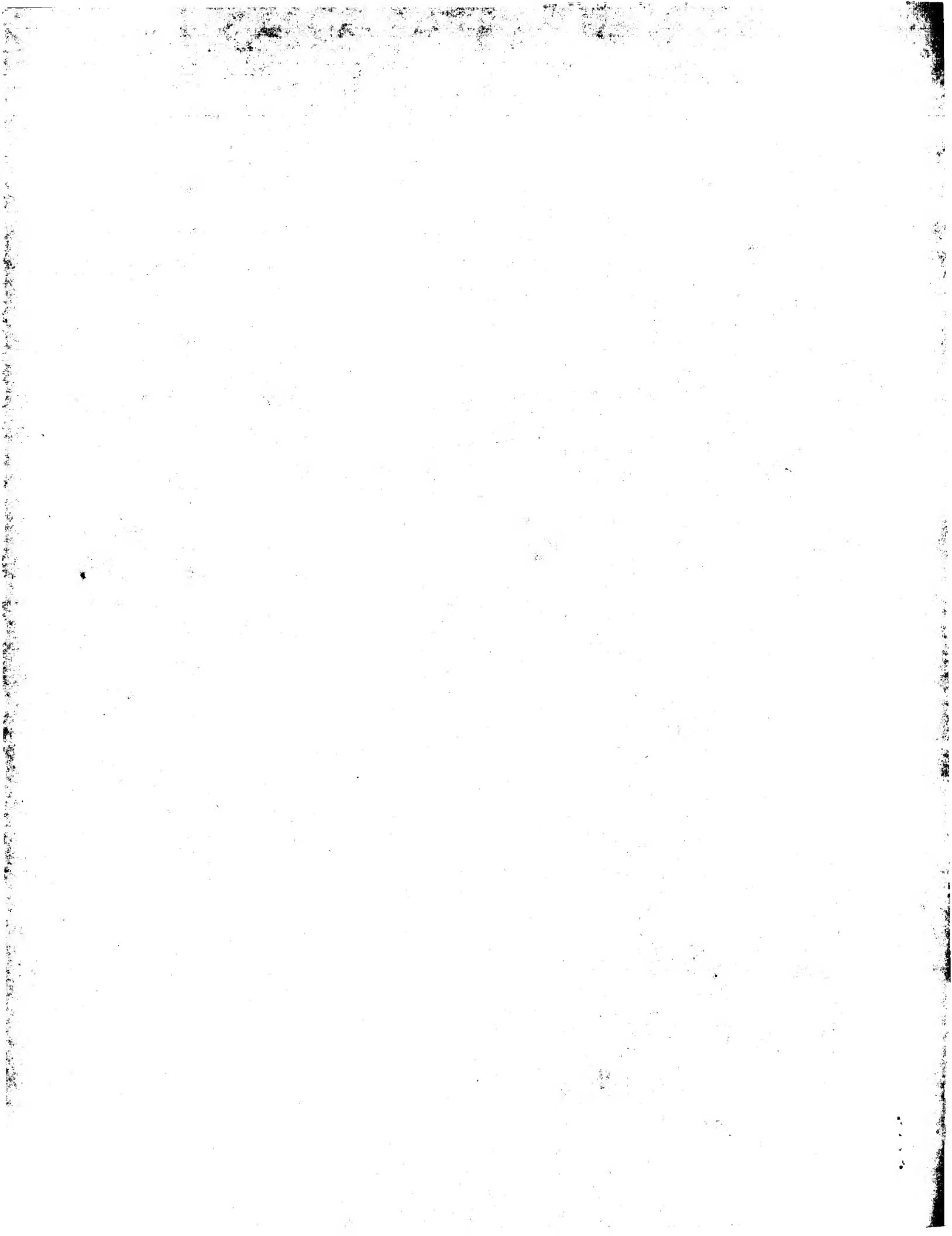
Query Match 88.2%; Score 30; DB 2; Length 424;
Best Local Similarity 85.7%; Pred. No. 67;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 369 LRLGTPA 375
|||||
RESULT 15
S15203
glycine hydroxymethyltransferase (EC 2.1.2.1) [validated] - Bradyrhizobium japonicum
C:Species: Bradyrhizobium japonicum
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 17-Mar-2000
C:Accession: S15203
R:Rosbach, S.; Hennecke, H. Mol. Microbiol. 5, 39-47, 1991
A:Title: Identification of glyA as a symbiotically essential gene in Bradyrhizobium j A:Reference number: S15203; MUID:91194557
A:Accession: S15203
A:Molecule type: DNA
A:Residues: 1-432 <ROS>
A:Cross-references: EMBL:X54638; NID:g39530; PIDN:CAA38450.1; PID:g39531
C:Genetics:
A:Gene: glyA
C:Function:
A:Description: EC 2.1.2.1 [validated; MUID:91194557]
A:Note: B. japonicum may have an additional pathway for glycine biosynthesis
C:Superfamily: glycine hydroxymethyltransferase
C:Keywords: phosphoprotein; pyridoxal phosphate; transferase
F:239/Active site: His #status predicted
F:240/Binding site: pyridoxal phosphate (Lys) (covalent) #status predicted

Query Match 88.2%; Score 30; DB 1; Length 432;
Best Local Similarity 85.7%; Pred. No. 68;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
 |||||
Db 371 LRLGTPA 377

Search completed: December 7, 2001, 00:26:05
Job time: 2845 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: December 7, 2001, 00:39:47 ; Search time 69.03 seconds
(without alignments)
3.718 Million cell updates/sec

Title: US-09-164-862B-3

Perfect score: 34

Sequence: 1 LRLGAPA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|--------------|---------------------|
| 1 | 34 | 100.0 | 383 | 1 C3LL_HUMAN | P36222 homo sapien |
| 2 | 34 | 100.0 | 504 | 1 GLYM_HUMAN | P34897 homo sapien |
| 3 | 34 | 100.0 | 504 | 1 GLYM_RABIT | P14519 oryctolagus |
| 4 | 31 | 91.2 | 427 | 1 GLYA_SYNY3 | P77962 synchocyst |
| 5 | 31 | 91.2 | 576 | 1 CWF_RHINE | P45404 rhizobium m |
| 6 | 30 | 88.2 | 224 | 1 RADP_PASMO | P57913 pasteurella |
| 7 | 30 | 88.2 | 322 | 1 KGP1_THETH | P21777 thermus aqu |
| 8 | 30 | 88.2 | 414 | 1 GLYA_BRAJA | P24060 bradyrhizob |
| 9 | 30 | 88.2 | 432 | 1 GLYC_SCHPO | Q10104 schizosacch |
| 10 | 30 | 88.2 | 472 | 1 GLYC_MOUSE | P50431 mus musculus |
| 11 | 30 | 88.2 | 478 | 1 GLYC_HUMAN | P34896 homo sapien |
| 12 | 30 | 88.2 | 483 | 1 GLYC_HUMAN | P07511 oryctolagus |
| 13 | 30 | 88.2 | 483 | 1 GLYC_RABIT | P35623 ovis aries |
| 14 | 30 | 88.2 | 483 | 1 GLYC_SHEEP | P75154 homo sapien |
| 15 | 30 | 88.2 | 756 | 1 Y665_HUMAN | O95095 mycobacteri |
| 16 | 30 | 88.2 | 758 | 1 CSTA_MYCTU | P27434 escherichia |
| 17 | 29 | 85.3 | 337 | 1 YFGA_ECOLI | P34895 hyphomicrob |
| 18 | 29 | 85.3 | 434 | 1 GLYA_HYPME | P37291 saccharomyc |
| 19 | 29 | 85.3 | 469 | 1 GLYC_YEAST | O13426 candida alb |
| 20 | 29 | 85.3 | 470 | 1 GLYC_CANAL | P05539 rattus norv |
| 21 | 28 | 82.4 | 122 | 1 CAL2_RAT | O15499 homo sapien |
| 22 | 28 | 82.4 | 205 | 1 GSCL_HUMAN | P11037 orygia pseu |
| 23 | 28 | 82.4 | 230 | 1 VP26_NPVOP | P08358 autographa |
| 24 | 28 | 82.4 | 240 | 1 VP26_NPVAC | P50434 corynebacte |
| 25 | 28 | 82.4 | 260 | 1 GLYA_COR51 | P18407 neurospora |
| 26 | 28 | 82.4 | 354 | 1 ALC_NEUCR | O9rku4 streptomyce |
| 27 | 28 | 82.4 | 376 | 1 ALIC_STRCO | O86583 streptomyce |
| 28 | 28 | 82.4 | 376 | 1 TRMU_STRCO | P08873 tritirachiu |
| 29 | 28 | 82.4 | 384 | 1 PRKX_TRIAL | P47634 mycoplasma |
| 30 | 28 | 82.4 | 406 | 1 GLYA_MYCGL | O86565 streptomyce |
| 31 | 28 | 82.4 | 420 | 1 GLYA_STRCO | O66990 aquifex aeo |
| 32 | 28 | 82.4 | 422 | 1 PYRC_AQUAE | O53615 mycobacteri |
| 33 | 28 | 82.4 | 425 | 1 GLA2_MYCTU | |

RESULT 1

| ID | C3LL_HUMAN | STANDARD; | PRT; | 383 AA. |
|----|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------|------|---------|
| AC | P36222 | P30923; | | |
| DT | 01-JUL-1993 | (Rel. 26. Created) | | |
| DT | 01-JUN-1994 | (Rel. 29. Last sequence update) | | |
| DT | 20-AUG-2001 | (Rel. 40. Last annotation update) | | |
| DE | CHITINASE-3 LIKE PROTEIN 1 PRECURSOR (CARTILAGE GLYCOPROTEIN-39) (GP-39) | (39 KDA SYNIOVAL PROTEIN) (YKL-40). | | |
| DE | CHI3L1. | | | |
| GN | Homo sapiens (Human). | | | |
| OS | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A., AND PARTIAL SEQUENCE. | | | |
| RC | TISSUE=Cartilage; | | | |
| RX | MEDLINE=94064658; PubMed=8245017; | | | |
| RA | Hakala B.E., White C., Recklies A.D.; | | | |
| RT | "Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family."; | | | |
| RT | J. Biol. Chem. 268:25803-25810(1993). | | | |
| RL | [2] | | | |
| RN | SEQUENCE FROM N.A. | | | |
| RP | TISSUE=Blood; | | | |
| RC | MEDLINE=97386591; PubMed=9244440; | | | |
| RX | Rehli M., Krause S.W., Andresen R.; | | | |
| RA | "Molecular characterization of the gene for human cartilage gp-39 (CHI3L1), a member of the chitinase protein family and marker for late stages of macrophage differentiation."; | | | |
| RT | Genomics 43:221-225(1997). | | | |
| RL | [3] | | | |
| RN | SEQUENCE OF 22-45. | | | |
| RP | MEDLINE=90328983; PubMed=2375755; | | | |
| RX | Nyrkos P., Golds E.E.; | | | |
| RA | "Human synovial cells secrete a 39 kDa protein similar to a bovine mammary protein expressed during the non-lactating period."; | | | |
| RT | Biochem. J. 269:265-268(1990). | | | |
| RL | CC | | | |
| CC | - FUNCTION: MAY PLAY AN IMPORTANT ROLE IN THE CAPACITY OF CELLS TO RESPOND TO AND COPE WITH CHANGES IN THEIR ENVIRONMENT. | | | |
| CC | - SUBCELLULAR LOCATION: EXTRACELLULAR. | | | |
| CC | - TISSUE SPECIFICITY: PRESENT IN ARTICULAR CHONDROCYTES, SYNIOVAL CELLS AS WELL AS IN LIVER. UNDETECTABLE IN MUSCLE TISSUES, LUNG, PANCREAS, MONONUCLEAR CELLS, OR FIBROBLASTS. | | | |
| CC | - PATTERN: GLYCOSYLATED. | | | |
| CC | - SIMILARITY: BELONGS TO FAMILY 18 OF GLYCOSYL HYDROLASES. | | | |
| CC | ----- | | | |
| CC | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch). | | | |
| CC | ----- | | | |

053441 mycobacteri
Q9x794 mycobacteri
P50435 methylobact
P34898 neurospora
P50432 caenorhabdi
P37292 saccharomyc
Q9p1w0 chlamydia m
Q92831 chlamydia p
O84439 chlamydia t
O83349 treponema p
P49053 curvularia
Q60560 mesocricetu

```

DR EMBL; M80927; AAA16074.1; -.
DR EMBL; Y08374; CAA69661.1; -.
DR EMBL; Y08375; CAA69661.1; JOINED.
DR EMBL; Y08376; CAA69661.1; JOINED.
DR EMBL; Y08377; CAA69661.1; JOINED.
DR EMBL; Y08378; CAA69661.1; JOINED.
DR EMBL; Y08379; CAA69661.1; JOINED.
DR PIR; S10677; S10677.
DR PIR; A33162; A33162.
DR MIM; G01525; -.
DR InterPro; IPR001579; Chitinase_2.
DR InterPro; IPR001223; Glyco_hydro.18.
DR Pfam; PF00704; Glyco_hydro.18; 1.
DR PROSITE; PS01095; CHITINASE_18; FALSE_NEG.
KW Glycoprotein; Signal.
FT SIGNAL 1 21
FT CHAIN 22 383 CHITINASE-3 LIKE PROTEIN 1.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 383 AA; 42613 MW; 76ADD8298EEC2D1 CRC64;

Query Match 100.0%; Score 34; DB 1; Length 383;
Best Local Similarity 100.0%; Pred. No. 4.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
   |||||
Db 245 LRLGAPA 251

RESULT 2
GLYM_HUMAN STANDARD; PRT; 504 AA.
ID GLYM_HUMAN
AC P34897; O00740;
DT 01-FEB-1994 (Rel. 28, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SERINE HYDROXYMETHYLTRANSFERASE, MITOCHONDRIAL PRECURSOR (EC 2.1.2.1)
DE (SERINE METHYLASE) (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).
GN SHMT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 22-504 FROM N.A.
RX MEDLINE=97153149; PubMed=8999870;
RA Stover P.J., Chen L.H., Suh J.R., Stover D.M., Keyomarsi K., Shane B.;
RT "Molecular cloning, characterization, and regulation of the human
RT mitochondrial serine hydroxymethyltransferase gene.";
RL J. Biol. Chem. 272:1842-1848(1997).
RN [2]
RP SEQUENCE OF 31-504 FROM N.A.
RX MEDLINE=93280158; PubMed=8505317;
RA Garrow T.A., Brenner A.A., Whitehead M.V., Chen X.N., Duncan R.G.,
RA Korenberg J.R., Shane B.;
RT "Cloning of human cDNAs encoding mitochondrial and cytosolic serine
RT hydroxymethyltransferases and chromosomal localization.";
RL J. Biol. Chem. 268:11910-11916(1993).
RN [3]
RP SEQUENCE OF 1-35 FROM N.A.
RX MEDLINE=20290786; PubMed=10828359;
RA Snell K., Baumann U., Byrne P.C., Chave K.J., Renwick S.B.,
RA Sanders P.G., Whitehouse S.K.;
RT "The genetic organization and protein crystallographic structure of
RT human serine hydroxymethyltransferase.";
RL Adv. Enzyme Regul. 40:353-403(2000).
CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC HORMONES AND OTHER COMPONENTS.
CC -1- SUBUNIT: HOMOTETRAMER.

-1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
-1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO FORMS OF THE ENZYMES: A
CYTOSOLIC ONE AND A MITOCHONDRIAL ONE.
-1- SIMILARITY: BELONGS TO THE SHMT FAMILY.

-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; U23143; AAA64572.1; -.
EMBL; L11932; AAA63258.1; -.
EMBL; Y12331; CAA72999.1; -.
PIR; B46746; B46746.
MIM; 138450; -.
InterPro; IPR001085; SHMT.
Pfam; PF00464; SHMT; 1.
PROSITE; PS00096; SHMT; 1.
Transferrase; Methyltransferase; Pyridoxal phosphate;
One-carbon metabolism; Mitochondrion; Transit peptide.
TRANSIT 1 29 MITOCHONDRION.
CHAIN 30 504 SERINE HYDROXYMETHYLTRANSFERASE.
BINDING 280 280 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CONFLICT 308 308 P -> L (IN REF. 2).
FT SEQUENCE 504 AA; 55992 MW; 7A13AF741C68FFD6 CRC64;
SQ

Query Match 100.0%; Score 34; DB 1; Length 504;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
   |||||
Db 424 LRLGAPA 430

RESULT 3
GLYM_RABIT STANDARD; PRT; 504 AA.
ID GLYM_RABIT
AC P14519; P79219;
DT 01-JAN-1990 (Rel. 13, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 13-JUL-1999 (Rel. 38, Last annotation update)
DE SERINE HYDROXYMETHYLTRANSFERASE, MITOCHONDRIAL PRECURSOR (EC 2.1.2.1)
DE (SERINE METHYLASE) (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).
GN SHMT2.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=NEW ZEALAND WHITE; TISSUE=Liver;
RA Whitehouse S.K., Sanders P.G., Snell K.;
RA Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 30-504.
RX TISSUE=Liver;
RX MEDLINE=89255303; PubMed=2656682;
RA Martini F., Maras B., Tanci P., Angelaccio S., Pascarella S.,
RA Barra D., Bossa F., Schirch V.;
RT "The primary structure of rabbit liver mitochondrial serine
RT hydroxymethyltransferase.";
RL J. Biol. Chem. 264:8509-8519(1989).
CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC HORMONES AND OTHER COMPONENTS.

```

CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL.
CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO FORMS OF THE ENZYMES: A
CC CYTOSOLIC ONE AND A MITOCHONDRIAL ONE.
CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X91902; CAA62998.1; -.
DR PIR: A33696; A33696.
DR InterPro: IPR001085; SHMT.
DR Pfam: PF00464; SHMT; 1.
DR PROSITE: PS00096; SHMT; 1.
KW Transferrase; Pyridoxal phosphate; One-carbon metabolism;
KW Mitochondrion; Transf. peptide.
FT TRANSIT 1 29 MITOCHONDRION.
FT CHAIN 30 504 SERINE HYDROXYMETHYLTRANSFERASE.
FT BINDING 280 280 PYRIDOXAL PHOSPHATE.
FT CONFLICT 81 81 I -> S (IN REF. 2).
FT CONFLICT 481 481 H -> R (IN REF. 2).
FT SEQUENCE 504 AA; 55901 MW; BE598B9CC0A1B13F CRC64;
SQ

Query Match 100.0%; Score 34; DB 1; Length 504;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 424 LRLGAPA 430
|||||

RESULT 4
GLYA_SVNY3 STANDARD; PRT; 427 AA.
AC P77962;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
DE (SHMT).
GN GLYA OR SLL1931.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
"Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res 3:109-136(1996).
CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC HORMONES AND OTHER COMPONENTS.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D90903; BAA17124.1; -.
DR InterPro: IPR001085; SHMT.
DR Pfam: PF00464; SHMT; 1.
DR PROSITE: PS00096; SHMT; 1.
KW Transferrase; Pyridoxal phosphate; One-carbon metabolism;
KW Complete proteome.
FT BINDING 231 231 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT SEQUENCE 427 AA; 46259 MW; B88AE834E2FA6045 CRC64;
SQ

Query Match 91.2%; Score 31; DB 1; Length 427;
Best Local Similarity 85.7%; Pred. No. 22;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
Db 362 LRLGSPA 368
|||||

RESULT 5
CCMF_RHIME STANDARD; PRT; 676 AA.
ID CCMF_RHIME
AC P45404;
DT 01-NOV-1995 (Rel. 32, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE CYTOCHROME C-TYPE BIOGENESIS PROTEIN CYCK.
GN CYCK.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=582;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAINE-AK631;
RX MEDLINE=95231514; PubMed=7715602;
RA Kereszt A., Slaska-Kiss K., Putnoky P., Banfalvi Z., Kondorosi A.;
RT "The cycHJKL genes of Rhizobium meliloti involved in cytochrome c
RT biogenesis are required for 'respiratory' nitrate reduction ex planta
RT and for nitrogen fixation during symbiosis.";
RL Mol. Gen. Genet. 247:39-47(1995).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RA Galibert F., Capela D., Hubler-Barloy F., Gatiou M., Batut J.,
RA Boistard P., Gouzy J., Kahn D., Thebaud P., Goffeau A.,
RA Purnelle B., Pohl T., Bothe G., Schneider S., Portetel D.,
RA Vandenbol M., Puchler A., Becker A., Weidner S.;
RL Submitted (MAR-2000) to the SWISS-PROT data bank.
CC -1- FUNCTION: REQUIRED FOR THE BIOGENESIS OF C-TYPE CYTOCHROMES.
CC -1- POSSIBLE SUBUNIT OF A HEME LYASE.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE
CC (POTENTIAL).
CC -1- SIMILARITY: BELONGS TO THE CCMF/CYCK/CCL1/NREE/CCSA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X82560; CAA57906.1; -.
DR InterPro: IPR002541; CytC_asm.

Tue Dec 11 08:46:50 2001

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DR InterPro: IPR003567; Cyt_c_biog.
 DR Pfam: PF01578; CytC_asm.1;
 KW Cytochrome c-type biogenesis; Transmembrane; Inner membrane.
 FT TRANSMEM 8 28 POTENTIAL.
 FT TRANSMEM 42 62 POTENTIAL.
 FT TRANSMEM 94 114 POTENTIAL.
 FT TRANSMEM 123 143 POTENTIAL.
 FT TRANSMEM 175 195 POTENTIAL.
 FT TRANSMEM 210 230 POTENTIAL.
 FT TRANSMEM 233 253 POTENTIAL.
 FT TRANSMEM 273 293 POTENTIAL.
 FT TRANSMEM 311 331 POTENTIAL.
 FT TRANSMEM 356 376 POTENTIAL.
 FT TRANSMEM 393 413 POTENTIAL.
 FT TRANSMEM 418 438 POTENTIAL.
 FT TRANSMEM 445 465 POTENTIAL.
 FT TRANSMEM 493 513 POTENTIAL.
 FT TRANSMEM 617 637 POTENTIAL.
 FT TRANSMEM 319 319 V -> I (IN REF. 1).
 FT CONFLICT 658 658 H -> R (IN REF. 1).
 FT CONFLICT 658 658
 SQ SEQUENCE 676 AA; 72610 MW; A2DD652280895E0D CRC64;

Query Match 91.2%; Score 31; DB 1; Length 676;
 Best Local Similarity 85.7%; Pred. No. 35;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
 |||||
 DB 642 LRVGAPA 648

RESULT 6
 RADC_PASMU STANDARD; PRT; 224 AA.
 ID RADC_PASMU
 AC P57913;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE DNA REPAIR PROTEIN RADC HOMOLOG.
 GN RADC OR PM1152.
 OS Pasteurella multocida.
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
 OC Pasteurella.
 OX NCBI_TaxID=747;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PM70;
 RX MEDLINE=21145866; PubMed=11248100;
 RA May B.J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;
 RT "Complete genomic sequence of Pasteurella multocida Pm70.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
 CC -1- FUNCTION: INVOLVED IN DNA REPAIR (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE RADC FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: AE006156; AAK03236.1;
 DR InterPro: IPR001405; RADC.
 DR ProDom: PD007415; RADC; 1.
 DR PROSITE: PS01302; RADC; 1.
 KW DNA repair; Complete proteome.
 SQ SEQUENCE 224 AA; 25541 MW; F8E9C447F1AD4FF6 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 224;
 Best Local Similarity 85.7%; Pred. No. 19;

QY 1 LRLGAPA 7
 |||||
 DB 17 LRVGAPA 23

RESULT 7
 K6P1_THETH STANDARD; PRT; 322 AA.
 ID K6P1_THETH
 AC P21777;
 DT 01-MAY-1991 (Rel. 18, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE 6-PHOSPHOFRUCTOKINASE ISOZYME 1 (EC 2.7.1.11) (PHOSPHOFRUCTOKINASE-1)
 DE (PFK1).
 OS Thermus aquaticus (subsp. thermophilus).
 OC Bacteria; Thermus/Delinococcus group; Thermus group; Thermus.
 OX NCBI_TaxID=274;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=HB8 / ATCC 27634;
 RX MEDLINE=91248220; PubMed=1828151;
 RA Xu J., Seki M., Denda K., Yoshida M.;
 RT "Molecular cloning of phosphofructokinase 1 gene from a thermophilic
 bacterium, Thermus thermophilus.";
 RL Biochem. Biophys. Res. Commun. 176:1313-1318(1991).
 RN [2]
 RP SEQUENCE OF 1-25.
 RC STRAIN=HB8 / ATCC 27634;
 RX MEDLINE=91324297; PubMed=1830879;
 RA Xu J., Oshima T., Yoshida M.;
 RT "Phosphoenolpyruvate-insensitive phosphofructokinase isozyme from
 Thermus thermophilus HB8.";
 RL J. Biochem. 109:199-203(1991).
 CC -1- CATALYTIC ACTIVITY: ATP + D-FRUCTOSE 6-PHOSPHATE = ADP +
 CC D-FRUCTOSE 1,6-BISPHOSPHATE.
 CC -1- ENZYME REGULATION: ALLOSTERICALLY INHIBITED BY PHOSPHOENOLPYRUVATE
 CC WHICH INDUCES THE DISSOCIATION OF THE ACTIVE TETRAMER INTO AN
 CC INACTIVE TWO-SUBUNIT FORMS.
 CC -1- PATHWAY: KEY CONTROL STEP OF GLYCOLYSIS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE PHOSPHOFRUCTOKINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: M71213; AAA27501.1;
 DR PIR: JQ1028; JQ1028.
 DR HSSP: P00512; 6PFK.
 DR InterPro: IPR000023; Phosphofructokinase.
 DR Pfam: PF00365; PFK; 1.
 DR PRINTS: PR00476; PFRCTKINASE.
 DR ProDom: PD000707; Phosphofructokinase; 1.
 DR PROSITE: PS00433; PHOSPHOFRUCTOKINASE; 1.
 KW Kinase; Transferase; Glycolysis; Allosteric enzyme; Multigene family.
 SQ SEQUENCE 322 AA; 33606 MW; 92945F3B204A9D15 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 322;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 LRLGAPA 7
 |||||
 DB 269 LRVGAPA 274

```

RESULT 8
GLYA_CAMJE
ID GLYA_CAMJE STANDARD; PRT; 414 AA.
AC P24531; Q9PIA3;
DT 01-MAR-1992 (Rel. 21, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
DE (SHMT).
GN GLYA OR CJO402.
OS Campylobacter jejuni.
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;
OC Campylobacter.
OX NCBI_TaxID=197;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 43431 / TGH 9011;
RX MEDLINE=91285434; PubMed=2060796;
RA Chan V.L., Bingham H.L.;
RT "Complete sequence of the Campylobacter jejuni glyA gene encoding
RT serine hydroxymethyltransferase."
RL Gene 101:51-58(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NCTC 11168;
RX MEDLINE=20150912; PubMed=10688204;
RA Parkhill J., Wren B.W., Mungall K., Kettle J.M., Churcher C.,
RA Basham D., Chillingworth T., Davies R.M., Feltwell T., Holtroyd S.,
RA Jagels K., Karlyshev A.V., Moule S., Pallen M.J., Penn C.W.,
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,
RA Whitehead S., Barrall B.G.;
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni
RT reveals hypervariable sequences."
RL Nature 403:665-668(2000).
CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC HORMONES AND OTHER COMPONENTS.
CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X53816; CAA37812.1; -.
DR EMBL; AL139075; CAB74238.1; -.
DR PIR; J01016; J01016.
DR PIR; S16468; S16468.
DR InterPro; IPR001085; SHMT.
DR Pfam; PF00464; SHMT; 1.
DR PROSITE; PS00096; SHMT; 1.
KW Transferase; Pyridoxal phosphate; One-carbon metabolism;
KW Complete proteome.
FT BINDING 224 224 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT CONFLICT 188 188 I -> V (IN REF. 1).
FT CONFLICT 214 214 Y -> H (IN REF. 1).
FT CONFLICT 386 386 V -> I (IN REF. 1).
SQ SEQUENCE 414 AA; 45777 MW; 91225A45CD9193E6 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 414;
Best Local Similarity 85.7%; Pred. No. 35;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 LRLGAPA 7
DB 355 LRLGTPA 361

RESULT 9
GLYA_BRAJA
ID GLYA_BRAJA STANDARD; PRT; 432 AA.
AC P24060;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
DE (SHMT).
GN GLYA.
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Bradyrhizobium group; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA 110SPC4;
RX MEDLINE=91194557; PubMed=2014004;
RA Rossbach S., Hennecke H.;
RT "Identification of glyA as a symbiotically essential gene in
RT Bradyrhizobium japonicum."
RL Mol. Microbiol. 5:39-47(1991).
CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE. ESSENTIAL FOR
CC EFFECTIVE NITROGEN-FIXING ROOT NODULE SYMBIOSIS.
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC HORMONES AND OTHER COMPONENTS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X54638; CAA38450.1; -.
DR PIR; S15203; S15203.
DR InterPro; IPR001085; SHMT.
DR Pfam; PF00464; SHMT; 1.
DR PROSITE; PS00096; SHMT; 1.
KW Transferase; Pyridoxal phosphate; One-carbon metabolism;
KW Nitrogen fixation.
FT BINDING 240 240 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
SQ SEQUENCE 432 AA; 45991 MW; E7DA7AF7D448307E CRC64;

Query Match 88.2%; Score 30; DB 1; Length 432;
Best Local Similarity 85.7%; Pred. No. 37;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

OY 1 LRLGAPA 7
DB 371 LRLGTPA 377

RESULT 10
GLYC_SCHPO
ID GLYC_SCHPO STANDARD; PRT; 472 AA.
AC Q10104;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE PROBABLE SERINE HYDROXYMETHYLTRANSFERASE, CYTOSOLIC (EC 2.1.2.1)

```

(SERINE METHYLASE) (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).
SPAC1866.04C.
Schizosaccharomyces pombe (Fission yeast).
Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
Schizosaccharomycetales; Schizosaccharomycetaceae;
Schizosaccharomycetes.
NCBI_TaxID=4896;
[1]
SEQUENCE FROM N.A.
STRAIN=972;
Connor R., Churcher C.M., Barrell B.G., Rajandream M.A., Walsh S.V.;
Submitted (DEC-1995) to the EMBL/GenBank/DDJB databases.
-1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE (BY SIMILARITY).
-1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
H(2)O = TETRAHYDROFOLATE + L-SERINE.
-1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
-1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
HORMONES AND OTHER COMPONENTS.
-1- SUBUNIT: HOMOTETRAMER.
-1- SUBCELLULAR LOCATION: CYTOPLASMIC.
-1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO FORMS OF THE ENZYMES: A
CYTOSOLIC ONE AND A MITOCHONDRIAL ONE.
-1- SIMILARITY: BELONGS TO THE SHMT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF237702; AAK15040.1; -
EMBL; X94479; CAA64226.1; -
EMBL; X94478; CAA64225.1; -
PDB; 1EJI; 03-NOV-00.
DR SWISS-2DPAGE; P50431; MOUSE.
DR MGI; 98299; Shmt1.
DR InterPro: IPR001085; SHMT.
DR Pfam: PF00464; SHMT; 1.
DR PROSITE; PS00096; SHMT; 1.
DR TRANSFERASE; Pyridoxal phosphate; One-carbon metabolism; 3D-structure.
KW BINDING 251 251
FT BINDING 249 249 PYRIDOXAL PHOSPHATE (BY SIMILARITY).
FT SEQUENCE 472 AA; 52080 MW; FEFD46161F9044FF CRC64;
SQ
Query Match 88.28; Score 30; DB 1; Length 472;
Best Local Similarity 85.7%; Pred. No. 40;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LRLGAPA 7
Db 392 LRLGTPA 398
|||||
RESULT 11
GLYC_MOUSE STANDARD; PRT; 478 AA.
ID P50431; Q64508;
AC 01-OCT-1996 (Rel. 34, Created)
DT 20-AUG-2001 (Rel. 40, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SERINE HYDROXYMETHYLTRANSFERASE, CYTOSOLIC (EC 2.1.2.1) (SERINE
DE METHYLASE) (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).
GN SHMT1 OR SHMT
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RC STRAIN=129/SV; TISSUE=Liver;
RX MEDLINE=20519026; PubMed=11063567;
RA Szebenyi D.M.E., Liu X., Krikunov I.A., Stover P.J., Thiel D.J.;
RT "Structure of a murine cytoplasmic serine hydroxymethyltransferase
RT quinonoid ternary complex: evidence for asymmetric obligate dimers.";
RL Biochemistry 39:13313-13323(2000).
RN [2]
RP SEQUENCE OF 1-316 FROM N.A.
RX MEDLINE=97017131; PubMed=8663732;
Nakshatri H., Bouillet P., Bhat-Nakshatri P., Chambon P.;
"Isolation of retinoic acid-repressed genes from P19 embryonal
carcinoma cells.";
Gene 174:79-84(1996).
-1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
-1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
H(2)O = TETRAHYDROFOLATE + L-SERINE.
-1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
-1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
HORMONES AND OTHER COMPONENTS.
-1- SUBUNIT: HOMOTETRAMER.
-1- SUBCELLULAR LOCATION: CYTOPLASMIC.
-1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO FORMS OF THE ENZYMES: A
CYTOSOLIC ONE AND A MITOCHONDRIAL ONE.
-1- SIMILARITY: BELONGS TO THE SHMT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF237702; AAK15040.1; -
EMBL; X94479; CAA64226.1; -
EMBL; X94478; CAA64225.1; -
PDB; 1EJI; 03-NOV-00.
DR SWISS-2DPAGE; P50431; MOUSE.
DR MGI; 98299; Shmt1.
DR InterPro: IPR001085; SHMT.
DR Pfam: PF00464; SHMT; 1.
DR PROSITE; PS00096; SHMT; 1.
DR TRANSFERASE; Pyridoxal phosphate; One-carbon metabolism; 3D-structure.
KW BINDING 251 251
FT BINDING 267 309
FT CONFLICT 304 304 A -> R (IN REF. 2).
FT SEQUENCE 478 AA; 52584 MW; FD03256EB62FDA1D CRC64;
SQ
Query Match 88.2%; Score 30; DB 1; Length 478;
Best Local Similarity 85.7%; Pred. No. 41;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 LRLGAPA 7
Db 395 LRLGTPA 401
|||||
RESULT 12
GLYC_HUMAN STANDARD; PRT; 483 AA.
ID P34896;
AC 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE SERINE HYDROXYMETHYLTRANSFERASE, CYTOSOLIC (EC 2.1.2.1) (SERINE
DE METHYLASE) (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).
GN SHMT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93280158; PubMed=8505317;
RA Garrow T.A., Brenner A.A., Whitehead M.V., Chen X.-N., Duncan R.G.,
RA Korenberg J.R., Shane B.;
RT "Cloning of human cDNAs encoding mitochondrial and cytosolic serine
RT hydroxymethyltransferases and chromosomal localization.";
RL J. Biol. Chem. 268:11910-11916(1993).

RN RP SEQUENCE FROM N.A.
 RC TISSUE=Liver;
 RA Xu L., Mangum J.H., Robertson D.L.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Breast;
 RX MEDLINE=97209776; PubMed=9056951;
 RA Chave K.J., Snell K., Sanders P.G.;
 RT "Isolation and characterisation of human genomic sequences encoding
 cytosolic serine hydroxymethyltransferase.";
 RL Biochem. Soc. Trans. 25:53-53(1997).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.65 ANGSTROMS) OF 11-480.
 RX MEDLINE=98428667; PubMed=9753690;
 RA Renwick S.B., Snell K., Baumann U.;
 RT "The crystal structure of human cytosolic serine
 hydroxymethyltransferase: a target for cancer chemotherapy.";
 RL Structure 6:1105-1116(1998).
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 H(2)O -> TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 HORMONES AND OTHER COMPONENTS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
 PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO FORMS OF THE ENZYMES: A
 CYTOSOLIC ONE AND A MITOCHONDRIAL ONE.
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; L11931; AAA63257.1; -;
 DR EMBL; L23928; AAA36020.1; -;
 DR EMBL; L23928; AAA36019.1; -;
 DR EMBL; L23928; AAA36018.1; -;
 DR EMBL; Y14485; CAB54838.1; -;
 DR PIR; A46746; A46746.
 DR PIR; B874; 16-AUG-99.
 DR MIM; L82144; -;
 DR InterPro; IPR001085; SHMT.
 DR Pfam; PF00464; SHMT; 1.
 DR PROSITE; PS00096; SHMT; 1.
 KW Transferase: Pyridoxal phosphate: One-carbon metabolism;
 FT BINDING 257 257 PYRIDOXAL PHOSPHATE.
 FT VARSPLIC 274 312 MISSING (IN ISOFORM 2).
 FT VARSPLIC 273 352 MISSING (IN ISOFORM 3).
 SQ SEQUENCE 483 AA; 53082 MW; 6CDD6CA06D017C19 CRC64;
 Query Match 88.2%; Score 30; DB 1; Length 483;
 Best Local Similarity 85.7%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LRLGAPA 7
 DB 401 LRLGTPA 407
 RESULT 13
 GLYC_RABIT
 ID GLYC_RABIT STANDARD; PRT; 483 AA.

AC P07511;
 DT 01-APR-1988 (Rel. 07, Created)
 DT 01-APR-1988 (Rel. 07, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE SERINE HYDROXYMETHYLTRANSFERASE, CYTOSOLIC (EC 2.1.2.1) (SERINE
 METHYLASE) (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).
 GN SHMT1.
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver;
 RX MEDLINE=92392263; PubMed=1381582;
 RA Byrne P.C., Sanders P.G., Snell K.;
 RT "Nucleotide sequence and expression of a cDNA encoding rabbit liver
 cytosolic serine hydroxymethyltransferase.";
 RL Biochem. J. 286:117-123(1992).
 RN [2]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=87194733; PubMed=3553178;
 RA Martini F., Angelaccio S., Pascarella S., Barra D., Bossa F.,
 RA Schirch V.;
 RT "The primary structure of rabbit liver cytosolic serine
 hydroxymethyltransferase.";
 RL J. Biol. Chem. 262:5499-5509(1987).
 RN [3]
 RP SEQUENCE OF 1-14, AND DEAMINATION OF ASN-5.
 RX MEDLINE=90202954; PubMed=2318867;
 RA Artigues A., Birkett A., Schirch V.;
 RT "Evidence for the in vivo deamidation and isomerization of an
 asparaginyl residue in cytosolic serine hydroxymethyltransferase.";
 RL J. Biol. Chem. 265:4853-4858(1990).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.8 ANGSTROMS).
 RC STRAIN=NEW ZEALAND WHITE; TISSUE=Liver;
 RX MEDLINE=93315211; PubMed=10387080;
 RA Scarsdale J.N., Kazanina G., Radaev S., Schirch V., Wright H.T.;
 RT "Crystal structure of rabbit cytosolic serine hydroxymethyltransferase
 at 2.8-A resolution: mechanistic implications.";
 RL Biochemistry 38:8347-8358(1999).
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 H(2)O -> TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 HORMONES AND OTHER COMPONENTS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO FORMS OF THE ENZYMES: A
 CYTOSOLIC ONE AND A MITOCHONDRIAL ONE.
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Z11846; CAA77870.1; -;
 DR PIR; A29140; XYRBSG.
 DR PIR; A35555; A35555.
 DR PIR; B35555; B35555.
 DR PIR; S24342; S24342.
 DR PDB; 1CJ0; 06-MAY-99.
 DR InterPro; IPR001085; SHMT.
 DR Pfam; PF00464; SHMT; 1.
 DR PROSITE; PS00096; SHMT; 1.
 KW Transferase: Pyridoxal phosphate; Acetylation: One-carbon metabolism;

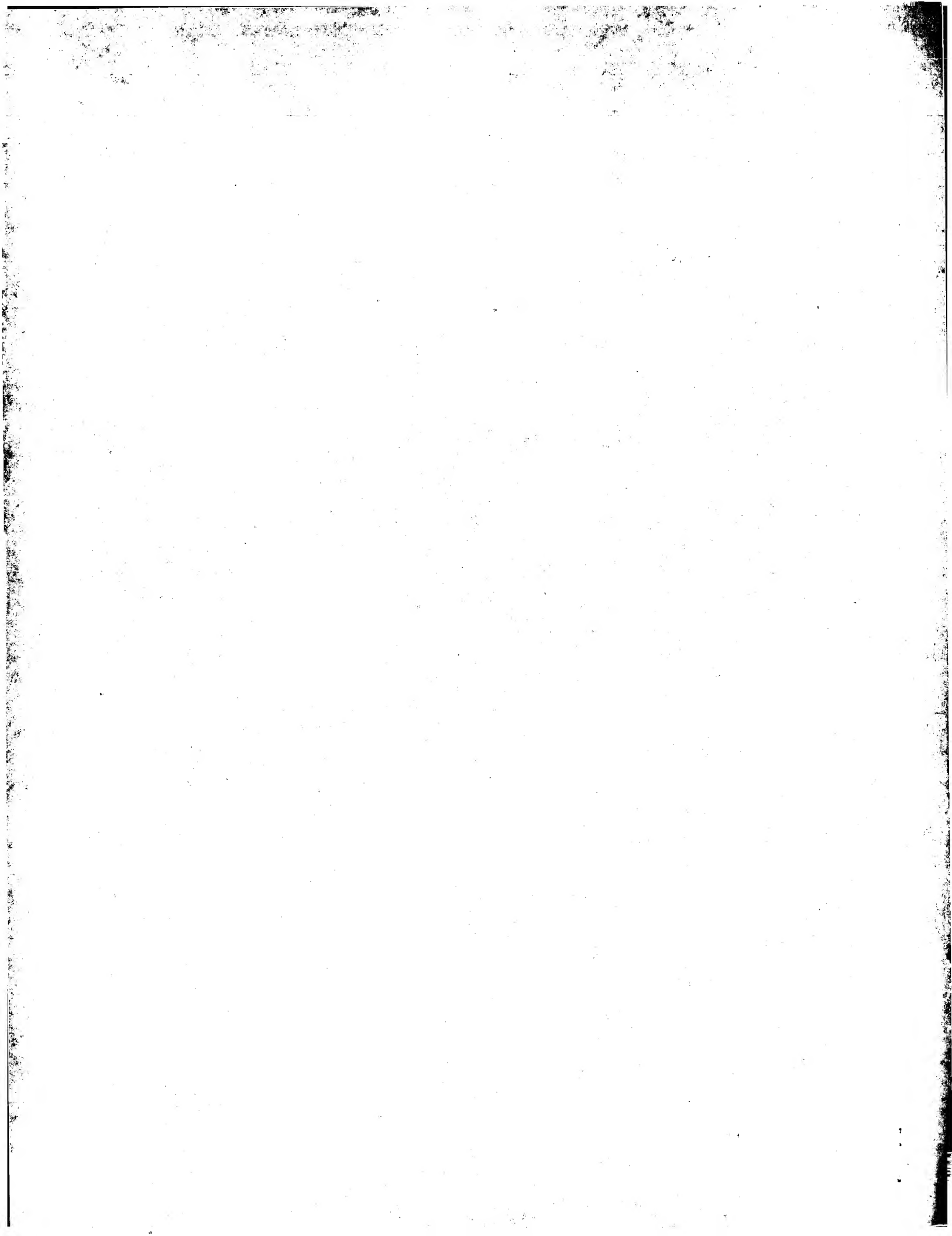
KW 3D-structure.
 FT INIT_MET 0 0
 FT MOD_RES 1 1 ACETYLATION.
 FT MOD_RES 5 5 DEMILATION.
 FT BINDING 256 PYRIDOXAL PHOSPHATE.
 FT SEQUENCE 483 AA; 52844 MW; 2DF5D2549DF3E052 CRC64;
 SQ SEQUENCE 483 AA; 52844 MW; 2DF5D2549DF3E052 CRC64;
 Query Match 88.2%; Score 30; DB 1; Length 483;
 Best Local Similarity 85.7%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LRLGAPA 7
 DB 400 LRLGTPA 406
 RESULT 14
 GLYC_SHEEP STANDARD; PRT; 483 AA.
 AC P35623;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE SERINE HYDROXYMETHYLTRANSFERASE, CYTOSOLIC (EC 2.1.2.1) (SERINE
 METHYLASE) (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).
 GN SHMT1.
 OS Ovis aries (Sheep).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE.
 RC TISSUE=Liver;
 RX MEDLINE=9413764; PubMed=8305478;
 RA Usha R., Savithri H.S., Rao N.A.;
 RT "The primary structure of sheep liver cytosolic serine
 hydroxymethyltransferase and an analysis of the evolutionary
 relationships among serine hydroxymethyltransferases.";
 RL Biochim. Biophys. Acta 1204:75-83(1994).
 RN [2]
 RP SEQUENCE FROM N.A., SEQUENCE OF 1-14, AND REVISIONS.
 RC TISSUE=Liver;
 RX MEDLINE=95331290; PubMed=7607226;
 RA Jagath-Reddy J., Ganesan K., Savithri H.S., Datta A., Rao N.A.;
 RT "cDNA cloning, overexpression in Escherichia coli, purification and
 characterization of sheep liver cytosolic serine
 hydroxymethyltransferase.";
 RL Eur. J. Biochem. 230:533-537(1995).
 RN [3]
 RP REVISIONS.
 RA Rao N.A.;
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE.
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 H(2)O = TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE.
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 HORMONES AND OTHER COMPONENTS.
 CC -1- SUBUNIT: HOMOTETRAMER.
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.
 CC -1- MISCELLANEOUS: IN EUKARYOTES THERE ARE TWO FORMS OF THE ENZYMES: A
 CYTOSOLIC ONE AND A MITOCHONDRIAL ONE.
 CC -1- SIMILARITY: BELONGS TO THE SHMT FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

CC EMBL; X80024; CAA56326.1; -.
 DR PIR; S40497; S40497.
 DR InterPro; IPR001085; SHMT.
 DR Pfam; PF00464; SHMT; 1.
 DR PROSITE; PS00096; SHMT; 1.
 KW Transferase; Pyridoxal phosphate; Acetylation; One-carbon metabolism.
 FT INIT_MET 0 0
 FT MOD_RES 1 1 ACETYLATION.
 FT BINDING 256 PYRIDOXAL PHOSPHATE.
 FT VARIANT 467 467 E -> V.
 FT VARIANT 473 473 T -> A.
 FT SEQUENCE 483 AA; 52893 MW; 812EAL913E729292 CRC64;
 SQ SEQUENCE 483 AA; 52893 MW; 812EAL913E729292 CRC64;
 Query Match 88.2%; Score 30; DB 1; Length 483;
 Best Local Similarity 85.7%; Pred. No. 41;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 LRLGAPA 7
 DB 400 LRLGTPA 406
 RESULT 15
 Y665_HUMAN STANDARD; PRT; 756 AA.
 ID Y665_HUMAN Q9H1G0; Q9H155;
 AC 075154; Q9NU10; Q9H1G0; Q9H155;
 DT 20-AUG-2001 (Rel. 40, Created)
 DT 20-AUG-2001 (Rel. 40, Last sequence update)
 DT 20-AUG-2001 (Rel. 40, Last annotation update)
 DE HYPOTHETICAL PROTEIN KIAA0665.
 GN KIAA0665.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=98403880; PubMed=9734811;
 RA Ishikawa K.-I., Nagase T., Suyama M., Miyajima N., Tanaka A.,
 RA Kotani H., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. X.
 The complete sequences of 100 new cDNA clones from brain which can
 code for large proteins in vitro.";
 RL DNA Res. 5:169-176(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wallis J., Lloyd C., Hall R.;
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 2 EF-HAND CALCIUM-BINDING DOMAINS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC EMBL; AB014565; BAA31640.1; -.
 DR EMBL; AL023881; CAB92745.1; -.
 DR EMBL; AL049542; CAC17519.1; -.
 DR EMBL; Z98882; CAC17523.1; -.
 DR InterPro; IPR002048; EF-hand.
 DR Pfam; PF00036; efhand; 2.
 KW Hypothetical protein; Calcium-binding; Repeat.
 FT CA_BIND 215 226 EF-HAND 1 (POTENTIAL).
 FT CA_BIND 247 258 EF-HAND 2 (POTENTIAL).
 SQ SEQUENCE 756 AA; 82439 MW; 264CEC399F28AFB9 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 756;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAP 6
Db 39 LRLGAP 44

Search completed: December 7, 2001, 00:39:48
Job time: 902 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 7, 2001, 00:38:26 ; Search time 135.48 Seconds
(without alignments)
7.358 Million cell updates/sec

Title: US-09-164-862B-3
Perfect score: 34
Sequence: 1 LRLGAPA 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_17:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | DB ID | Description |
|------------|-------|---------------|--------|-----------|---------------------|
| 1 | 34 | 100.0 | 332 | 6 O18949 | O18949 bos taurus |
| 2 | 34 | 100.0 | 383 | 6 Q29411 | Q29411 sus scrofa |
| 3 | 34 | 100.0 | 504 | 11 Q9CZN7 | Q9czn7 mus musculus |
| 4 | 34 | 100.0 | 504 | 11 Q99K87 | Q99k87 mus musculus |
| 5 | 31 | 91.2 | 117 | 10 Q9SU32 | Q9su32 arabidopsis |
| 6 | 31 | 91.2 | 535 | 2 Q9A575 | Q9a575 caulobacter |
| 7 | 31 | 91.2 | 664 | 2 Q52820 | Q52820 rhizobium l |
| 8 | 30 | 88.2 | 92 | 4 Q9UMC9 | Q9umc9 homo sapien |
| 9 | 30 | 88.2 | 150 | 6 Q9TSA5 | Q9tsa5 ovnis aries |
| 10 | 30 | 88.2 | 185 | 10 Q9AY89 | Q9ay89 oryza sativ |
| 11 | 30 | 88.2 | 229 | 2 Q60108 | Q60108 versinia en |
| 12 | 30 | 88.2 | 238 | 5 Q9VVF0 | Q9vvf0 drosophila |
| 13 | 30 | 88.2 | 310 | 4 Q9UMD2 | Q9umd2 homo sapien |
| 14 | 30 | 88.2 | 352 | 11 Q9WTV1 | Q9wtv1 rattus norv |
| 15 | 30 | 88.2 | 361 | 2 Q9X7G7 | Q9x7g7 methylobact |
| 16 | 30 | 88.2 | 386 | 4 Q75086 | Q75086 homo sapien |
| 17 | 30 | 88.2 | 390 | 4 Q75900 | Q75900 homo sapien |
| 18 | 30 | 88.2 | 390 | 4 Q9UBR9 | Q9ubr9 homo sapien |
| 19 | 30 | 88.2 | 398 | 2 Q9ZHQ7 | Q9zqh7 streptomyce |

| | | | | | |
|----|----|------|------|-----------|--------------------|
| 20 | 30 | 88.2 | 407 | 2 Q9RD17 | Q9rd17 streptomyce |
| 21 | 30 | 88.2 | 408 | 2 Q9Z4V4 | Q9z4v4 streptomyce |
| 22 | 30 | 88.2 | 423 | 1 Q9YB22 | Q9yb22 aeropyrum p |
| 23 | 30 | 88.2 | 424 | 2 Q9PET2 | Q9pet2 xyella fas |
| 24 | 30 | 88.2 | 432 | 2 Q9F5D2 | Q9f5d2 agrobacteri |
| 25 | 30 | 88.2 | 444 | 11 Q99KN1 | Q99kn1 mus musculu |
| 26 | 30 | 88.2 | 463 | 2 Q34474 | Q34474 bacillus su |
| 27 | 30 | 88.2 | 478 | 11 Q9CWR5 | Q9cwr5 mus musculu |
| 28 | 30 | 88.2 | 482 | 2 Q59326 | Q59326 clostridium |
| 29 | 30 | 88.2 | 485 | 3 O13920 | O13920 schizosacch |
| 30 | 30 | 88.2 | 486 | 2 O86746 | O86746 streptomyce |
| 31 | 30 | 88.2 | 514 | 1 Q34180 | Q34180 halobacteri |
| 32 | 30 | 88.2 | 533 | 2 Q9KBP1 | Q9kbp1 bacillus ha |
| 33 | 30 | 88.2 | 574 | 4 Q75134 | Q75134 homo sapien |
| 34 | 30 | 88.2 | 595 | 5 Q9W362 | Q9w362 drosophila |
| 35 | 30 | 88.2 | 663 | 2 Q915D9 | Q915d9 pseudomonas |
| 36 | 30 | 88.2 | 731 | 4 Q9BYJ6 | Q9byj6 homo sapien |
| 37 | 30 | 88.2 | 1016 | 2 Q59206 | Q59206 bradyrhizob |
| 38 | 30 | 88.2 | 1067 | 5 Q9GVH6 | Q9gvh6 caenorhabdi |
| 39 | 30 | 88.2 | 1224 | 8 Q35057 | Q35057 marchantia |
| 40 | 30 | 88.2 | 1239 | 2 Q9F824 | Q9fb24 streptomyce |
| 41 | 30 | 88.2 | 1248 | 4 Q95458 | Q95458 homo sapien |
| 42 | 30 | 88.2 | 1256 | 5 Q9VF50 | Q9vf50 drosophila |
| 43 | 30 | 88.2 | 1594 | 2 Q914N8 | Q914n8 pseudomonas |
| 44 | 30 | 88.2 | 1778 | 5 Q9VLD8 | Q9vld8 drosophila |
| 45 | 30 | 88.2 | 1326 | 4 Q9NQT8 | Q9nqt8 homo sapien |

ALIGNMENTS

RESULT 1
O18949 ID O18949 PRELIMINARY; PRT; 332 AA.
AC O18949;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE CHITINASE-LIKE PROTEIN 1 (FRAGMENT).
GN CLP-1.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Recklies A.D., White C.;
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF011373; AAB64304.1; -;
DR InterPro; IPR001223; Glyco_hydro_18.
DR Pfam; PF00704; Glyco_hydro_18; 1.
FT NON_TER 1
FT NON_TER 332
SQ SEQUENCE 332 AA; 37434 MW; 03F163323486408F CRC64;

Query Match 100.0%; Score 34; DB 6; Length 332;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7

Db 221 LRLGAPA 227

RESULT 2

Q29411 ID Q29411 PRELIMINARY; PRT; 383 AA.
AC Q29411;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)

Tue Dec 11 08:46:50 2001

```

DE 38 KDA HEPARIN-BINDING GLYCOPROTEIN.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1];
RP SEQUENCE FROM N.A.
RC TISSUE-SMOOTH MUSCLE; PubMed=7768902;
RX Shackleton L.M., Mann D.M., Millis A.J.;
RA "Identification of a 38-kDa heparin-binding glycoprotein (gp38k) in
RT differentiating vascular smooth muscle cells as a member of a group of
RT proteins associated with tissue remodeling.";
RL J. Biol. Chem. 270:13076-13083(1995).
RN [2];
RP SEQUENCE FROM N.A.
RC TISSUE-SMOOTH MUSCLE;
RA Shackleton L.M., Mann D.M., Millis A.J.T.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U19900; AAA86482.1; -.
DR EMBL; Z47803; CAA87764.1; -.
DR InterPro: IPR001223; Glyco_hydro_18.
DR Pfam: PF00704; Glyco_hydro_18; 1.
KW Heparin-binding.
SQ SEQUENCE 383 AA; 42443 MW; 3D1039F49910BDC4 CRC64;

Query Match 100.0%; Score 34; DB 6; Length 383;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
DB 245 LRLGAPA 251
|||||

RESULT 3
Q9CZN7 PRELIMINARY; PRT; 504 AA.
AC Q9CZN7;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (SERINE METHYLASE)
DE (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT).
GN 2700043D08R1K.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Harai A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saio R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Buit C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombarts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RA "Functional annotation of a full-length mouse cDNA collection.";

```

```

RL Nature 409:685-690(2001).
CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
CC HORMONES AND OTHER COMPONENTS.
CC -1- SIMILARITY: TO OTHER EUKARYOTIC, ARCHAEABACTERIAL AND PROKARYOTIC
CC MEMBERS OF THE SHMT FAMILY.
DR EMBL; AK012355; BAB28184.1; -.
DR MGD; MGI:1919660; 2700043D08R1K.
DR InterPro: IPR001085; SHMT.
DR Pfam: PF00464; SHMT; 1.
DR PROSITE; PS00096; SHMT; 1.
KW One-carbon metabolism; Pyridoxal phosphate; Transferase.
SQ SEQUENCE 504 AA; 55758 MW; 38A380996F75EC3C CRC64;

Query Match 100.0%; Score 34; DB 11; Length 504;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
DB 424 LRLGAPA 430
|||||

RESULT 4
Q99K87 PRELIMINARY; PRT; 504 AA.
AC Q99K87;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE RIKEN CDNA 2700043D08 GENE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC004825; AAH04825.1; -.
SQ SEQUENCE 504 AA; 55760 MW; C494CF5A7B972481 CRC64;

Query Match 100.0%; Score 34; DB 11; Length 504;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
DB 424 LRLGAPA 430
|||||

RESULT 5
Q9SU32 PRELIMINARY; PRT; 117 AA.
AC Q9SU32;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE AIRIA-LIKE PROTEIN.
GN TIP17.120 OR AT4G12530.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1];
RP SEQUENCE FROM N.A.
RA Bevan M., Peters S.A., van Staveren M., Dirkse W., Stiekema W.,

```

RA Bancroft I., Mewes H.W., Mayer K.F.X., Lemcke K., Schueller C.;
 RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Arabidopsis sequencing project;
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Peters S.A., van Staveren M., Dirkse W., Stiekema W., Mewes H.W.,
 Lemcke K., Mayer K.F.X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Arabidopsis sequencing project;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL049730; CAB41723.1; -;
 DR EMBL; AL161534; CAB78296.1; -;
 DR HSSP; P24337; LHYP.
 DR InterPro; IPR003612; AAI.
 DR InterPro; IPR000528; Plant_LTP.
 DR Pfam; PF00279; LTP; 1.
 DR SMART; SM00499; AAI; 1.
 SQ SEQUENCE 117 AA; 12365 MW; A3382D90B2594A17 CRC64;

Query Match 91.2%; Score 31; DB 10; Length 117;
 Best Local Similarity 85.7%; Pred. No. 36;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
 Db 53 LRLGAPA 59

RESULT 6
 Q9A575 PRELIMINARY; PRT; 535 AA.
 AC Q9A575;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE ABC TRANSPORTER, ATP-BINDING PROTEIN.
 GN CC2587.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=69394;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21173698; PubMed=11259647;
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
 DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 Kolonay J.F., Smit J., Craven M.B., Khouri J., Shetty J., Berry K.,
 Uitterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of *Caulobacter crescentus*."
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005927; AAK24557.1; -;
 DR TIGR; CC2587; -;
 KW ATP-binding; Complete proteome.
 SQ SEQUENCE 535 AA; 56322 MW; E383572801C4FFE0 CRC64;

Query Match 91.2%; Score 31; DB 2; Length 535;
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
 Db 394 VRLGAPA 400

RESULT 7
 Q52820 PRELIMINARY; PRT; 664 AA.
 AC Q52820;
 DT 01-NOV-1996 (TREMBLrel. 01, Created)
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE DNA FOR CYCH, CYCJ, CYCK AND CYCL GENES.
 GN CYCK.
 OS Rhizobium leguminosarum.
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=384;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=8401;
 RX MEDLINE=95394794; PubMed=7665469;
 RA Delgado M.J., Yeoman K.H., Wu G., Vargas C., Davies A., Poole R.K.,
 Johnston A.W.B., Downie J.A.;
 RA "Characterization of the *cycHJKL* genes involved in cytochrome c
 biogenesis and symbiotic nitrogen fixation in *Rhizobium*
leguminosarum."
 RT J. Bacteriol. 177:4927-4934(1995).
 RL EMBL; X89726; CAA61878.1; -;
 DR InterPro; IPR002541; CytC_asm.
 DR InterPro; IPR003567; CytC_biol.
 DR Pfam; PF01578; CytC_asm; 1.
 DR PRINTS; PRO1410; CCBIOGENESIS.
 SQ SEQUENCE 664 AA; 71955 MW; 6FE95E5919FD8BC CRC64;

Query Match 91.2%; Score 31; DB 2; Length 664;
 Best Local Similarity 85.7%; Pred. No. 2e+02;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
 Db 644 LRVGAPA 650

RESULT 8
 Q9UMC9 PRELIMINARY; PRT; 92 AA.
 AC Q9UMC9;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CYTOSOLIC SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.2.1) (FRAGMENT).
 GN SHMT.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chave K.J., Snell K., Sanders P.G.;
 RT "Isolation and characterisation of human genomic sequences encoding
 cytosolic serine hydroxymethyltransferase."
 RL Biochem. Soc. Trans. 25:53-53(1997).
 DR EMBL; Y14489; CAB54842.1; -;
 DR HSSP; P34896; 1BJA
 DR InterPro; IPR001085; SHMT.
 DR Pfam; PF00464; SHMT; 1.
 KW Transferase; Methyltransferase.
 FT NON_TER 1
 FT NON_TER 92 92
 SQ SEQUENCE 92 AA; 10222 MW; 8BBE599C78753694 CRC64;

Query Match 88.2%; Score 30; DB 4; Length 92;
 Best Local Similarity 85.7%; Pred. No. 46;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Tue Dec 11 08:46:50 2001

QY 1 LRLGAPA 7
 Db 10 LRLGTPA 16

RESULT 9
 Q9TSA5 PRELIMINARY; PRT; 150 AA.
 AC Q9TSA5; (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBlrel. 17, Last annotation update)
 DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.1.2.1) (FRAGMENTS).
 OS Ovis aries (Sheep).
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
 OC Bovidae; Caprinae; Ovis.
 OX NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=93013843; PubMed=1398711;
 RA Usha R., Savithri H.S., Rao N.A.;
 RT "Partial amino acid sequence of sheep liver serine
 hydroxymethyltransferase and comparison of peptide maps of the enzyme
 from human, ox livers and Escherichia coli."
 RT Indian J. Biochem. Biophys. 29:183-188(1992).
 RL HSSP: P34896; 1B74.
 DR InterPro: IPR001085; SHMT.
 DR Pfam: PF00464; SHMT; 1.
 FT NON_TER 1
 FT NON_CONS 19 20
 FT NON_CONS 33 34
 FT NON_CONS 42 43
 FT NON_CONS 53 54
 FT NON_CONS 87 88
 FT NON_CONS 95 96
 FT NON_CONS 101 102
 FT NON_CONS 110 111
 FT NON_CONS 116 117
 FT NON_CONS 130 131
 FT NON_CONS 134 135
 FT NON_CONS 144 145
 FT NON_TER 150 150
 SQ SEQUENCE 150 AA; 16373 MW; 8F174CF96712FC1E CRC64;

Query Match 88.2%; Score 30; DB 6; Length 150;
 Best Local Similarity 85.7%; Pred. No. 75;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAPA 7
 Db 120 LRLGTPA 126

RESULT 10
 Q9AY89 PRELIMINARY; PRT; 185 AA.
 AC Q9AY89;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DE HYPOTHETICAL 19.9 KDA PROTEIN.
 OS Oryza sativa (Rice).
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. NIPPONBARE;
 RA Buell C.R., Yuan Q., Moffat K.S., Hill J.N., Burr P.C., Hsiao J.,

RA Zismann V., Pai G., Bowman C.L., Fujii C.Y., VanAken S.E.,
 Bowman C.L., Craven B., Utterback T.R., Khalak H., Feldblyum T.V.,
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;
 RT "Oryza sativa chromosome 3 BAC OSJNBa0004B24 genomic sequence."
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC084319; AAG59666.1;
 KW Hypothetical protein.
 SQ SEQUENCE 185 AA; 19873 MW; C24B058C54B89DA2 CRC64;

Query Match 88.2%; Score 30; DB 10; Length 185;
 Best Local Similarity 100.0%; Pred. No. 93;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLGAPA 7
 Db 40 RLGAPA 45

RESULT 11
 Q60108 PRELIMINARY; PRT; 229 AA.
 AC Q60108;
 DT 01-NOV-1996 (TREMBlrel. 01, Created)
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
 DE RNA POLYMERASE SIGMA FACTOR.
 OS Versinia enterocolitica.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Versinia.
 OX NCBI_TaxID=630;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=8081;
 RX MEDLINE=96249698; PubMed=8830263;
 RA Kapral V., Olson J.W., Pepe J.C., Miller V.L., Minnich S.A.;
 RT "Temperature-dependent regulation of Versinia enterocolitica Class III
 flagellar genes.";
 RL Mol. Microbiol. 19:1061-1071(1996).
 CC -1- FUNCTION: THE SIGMA FACTOR IS AN INITIATION FACTOR THAT PROMOTES
 CC ATTACHMENT OF THE RNA POLYMERASE TO SPECIFIC INITIATION SITES AND
 CC THEN IS RELEASED.
 CC -1- SIMILARITY: TO THE SIGMA-70 FACTOR FAMILY.
 DR EMBL; L33466; AAB38522.1;
 DR InterPro: IPR000943; Sigma_70.
 DR Pfam: PF00140; sigma70; 1.
 DR PROSITE; PS00715; SIGMA70_1; 1.
 DR PROSITE; PS00716; SIGMA70_2; 1.
 KW DNA-binding; DNA-directed RNA polymerase; Sigma factor;
 KW Transcription regulation.
 SQ SEQUENCE 229 AA; 26459 MW; D16E1FF20C34233C CRC64;

Query Match 88.2%; Score 30; DB 2; Length 229;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RLGAPA 7
 Db 99 RLGAPA 104

RESULT 12
 Q9VVF0 PRELIMINARY; PRT; 238 AA.
 ID Q9VVF0
 AC Q9VVF0;
 DT 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE CG6485 PROTEIN.
 GN CG6485.
 OS Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter A.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003524; AAF49361.1;
 DR FlyBase: FBgn0036706; CG6485.
 DR InterPro: IPR002023; Complex1_24kd.
 DR Pfam: PF01257; complex1_24kd; 1.
 DR ProDom: PD003859; Complex1_24kd; 1.
 SQ SEQUENCE 238 AA; 26704 MW; 6FE5471D8A72A8E9 CRC64;

Query Match 88.2%; Score 30; DB 5; Length 238;
 Best Local Similarity 100.0%; Pred. No. 1.2e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAP 6
 DB 17 LRLGAP 22
 RESULT 13
 Q9UMD2
 ID Q9UMD2 PRELIMINARY; PRT; 310 AA.
 AC Q9UMD2;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE SERINE HYDROXYMETHYLTRANSFERASE (EC 2.1.1.2.1) (SERINE METHYLASE)
 DE (GLYCINE HYDROXYMETHYLTRANSFERASE) (SHMT) (FRAGMENT).
 GN SHMT.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BREAST;
 RA Chave K.J., Snell K., Sanders P.G.;
 RT "Isolation and characterisation of human genomic sequences encoding
 RT cytosolic serine hydroxymethyltransferase.";
 RL Biochem Soc Trans. 25:53-53(1997).
 CC -1- FUNCTION: INTERCONVERSION OF SERINE AND GLYCINE (BY SIMILARITY).
 CC -1- CATALYTIC ACTIVITY: 5,10-METHYLENETETRAHYDROFOLATE + GLYCINE +
 CC H(2)O = TETRAHYDROFOLATE + L-SERINE.
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).
 CC -1- PATHWAY: KEY ENZYME IN THE BIOSYNTHESIS OF PURINES, LIPIDS,
 CC HORMONES AND OTHER COMPONENTS.
 CC -1- SIMILARITY: TO OTHER EUKARYOTIC, ARCHAEABACTERIAL AND PROKARYOTIC
 CC MEMBERS OF THE SHMT FAMILY.
 DR EMBL: Y14486; CAB54839.1; -.
 DR HSSP: P34896; 1BJ4.
 DR InterPro: IPR001085; SHMT.
 DR Pfam: PF00464; SHMT; 1.
 DR PROSITE: PS00096; SHMT; 1.
 DR Methytransferase; One-carbon metabolism; Pyridoxal phosphate;
 KW Transferase.
 FT NON_TER 1 1
 SQ SEQUENCE 310 AA; 33788 MW; 7CE6099CD03C51C5 CRC64;

Query Match 88.2%; Score 30; DB 4; Length 310;
 Best Local Similarity 85.7%; Pred. No. 1.6e+02;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 LRLGAP 7
 DB 228 LRLGTPA 234
 RESULT 14
 Q9WTV1
 ID Q9WTV1 PRELIMINARY; PRT; 352 AA.
 AC Q9WTV1;
 DT 01-NOV-1999 (TrEMBLrel. 12, Created)
 DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
 DE GLYCOPROTEIN-39 PRECURSOR (FRAGMENT).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=LEWIS;
 RA Wendling U., Boots A.M.H., van Eden W.;
 RT "Cloning of the rat homologue of Human Cartilage glycoprotein-39 a
 RT potential autoantigen in arthritis.";
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF062038; AAD22610.1; -.
 DR HSSP: P07254; 1CTN.
 DR InterPro: IPR000677; 2S_Globulin.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam: PF00704; Glyco_hydro_18; 1.
 DR PRINTS: PR00551; 2SGLOBULIN.
 DR NON_TER 1 1
 FT NON_TER 352 352
 SQ SEQUENCE 352 AA; 39391 MW; CBDE991610AC936C CRC64;

Query Match 88.2%; Score 30; DB 11; Length 352;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LRLGAP 6

Tue Dec 11 08:46:50 2001

|||||
Db 224 LRLGAP 229

RESULT 15
Q9X7G7 PRELIMINARY; PRT; 361 AA.
AC Q9X7G7;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ORF361.
OS Methylobacterium sp. (strain CM4).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Methylobacterium group; Methylobacterium.
OX NCBI_TaxID=82543;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99218329; PubMed=10200311;
RA Vannelli T., Messmer M., Studer A., Vuilleumier S., Leisinger T.;
RT "A corrinoid-dependent catabolic pathway for growth of a
RT Methylobacterium strain with chloromethane."
RL Proc. Natl. Acad. Sci. U.S.A. 96:4615-4620(1999).
DR EMBL: AJ011317; CAB40740.1; -.
DR InterPro: IPR02748; CblD.
DR Pfam: PF01888; CblD; 1.
SQ SEQUENCE 361 AA; 37534 MW; 4913FE8B70B3B3F2 CRC64;

Query Match 88.2%; Score 30; DB 2; Length 361;
Best Local Similarity 85.7%; Pred. No. 1.8e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 LRLGAPA 7
|||
Db 86 LRFGAPA 92

Search completed: December 7, 2001, 00:38:28
Job time: 907 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2001, 21:23:25 ; Search time 2421.69 Seconds
(without alignments)
11451.407 Million cell updates/sec

Title: US-09-164-862B-4
Perfect score: 1681
Sequence: 1 ctaggtagctggcaccagga.....tacaagagtttaacagtgtg 1681

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1472140 seqs, 8248589755 residues

Total number of hits satisfying chosen parameters: 2944280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl :

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.ov.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.om.*
- 20: em.or.*
- 21: em.ov.*
- 22: em.pat.*
- 23: em.ph.*
- 24: em.pl.*
- 25: em.ro.*
- 26: em.sts.*
- 27: em.sy.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htgo_hum.*
- 31: em.htgo_inv.*
- 32: em.htgo_rod.*
- 33: em.htg_hum.*
- 34: em.htg_inv.*
- 35: em.htg_rod.*
- 36: em.htg_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

§

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|--------|-------|--------|----|-----------|--------------------|
| 1 | 1575.4 | 93.7 | 1741 | 9 | HUMHA3G | M80927 Human glyco |
| 2 | 1543.2 | 91.8 | 1801 | 9 | BC008568 | BC008568 Homo sapi |
| 3 | 1120.2 | 66.6 | 1149 | 6 | E01500 | E01500 CDNA encodi |
| 4 | 1115.4 | 66.4 | 1149 | 6 | E01501 | E01501 CDNA encodi |
| 5 | 976 | 58.1 | 1733 | 4 | SSGP38KD | 247803 S.scrofa 38 |
| 6 | 976 | 58.1 | 1733 | 4 | SSU19900 | U19900 Sus scrofa |
| 7 | 775.2 | 46.1 | 1641 | 10 | BC004734 | BC004734 Mus muscu |
| 8 | 775.2 | 46.1 | 1651 | 10 | BC003780 | BC003780 Mus muscu |
| 9 | 775.2 | 46.1 | 1664 | 10 | BC005611 | X93035 M.musculus |
| 10 | 751 | 44.7 | 1616 | 10 | MMBRP39 | AF011373 Bos tauru |
| 11 | 749 | 44.6 | 997 | 4 | AF011373 | E01502 CDNA encodi |
| 12 | 744.2 | 44.3 | 966 | 6 | E01502 | AF062038 Rattus no |
| 13 | 736 | 43.8 | 1056 | 10 | AF062038 | AJ251847 Homo sapi |
| 14 | 521.8 | 31.0 | 2729 | 9 | HSR251847 | Y08378 H.sapiens g |
| 15 | 521.8 | 31.0 | 2788 | 9 | HSY08378 | AL357133 Homo sapi |
| 16 | 494.6 | 29.4 | 160010 | 2 | AL357133 | AL359090 Homo sapi |
| 17 | 494.6 | 29.4 | 171058 | 2 | AL359090 | AL451082 Homo sapi |
| 18 | 494.6 | 29.4 | 200794 | 2 | AL451082 | U29615 Homo chito |
| 19 | 397.6 | 23.7 | 1633 | 9 | HSU29615 | U62662 Homo sapien |
| 20 | 397.6 | 23.7 | 1710 | 9 | HSU62662 | AX108750 Sequence |
| 21 | 397.2 | 23.6 | 1636 | 6 | AX108750 | AX108752 Sequence |
| 22 | 396 | 23.6 | 1656 | 6 | AX108752 | U49835 Human YKL-3 |
| 23 | 384.6 | 22.9 | 1418 | 9 | HSU49835 | AR042834 Sequence |
| 24 | 384.6 | 22.9 | 1433 | 6 | AR042834 | U58514 Human chit1 |
| 25 | 384.6 | 22.9 | 1434 | 9 | HSU58514 | BC011460 Homo sapi |
| 26 | 383 | 22.8 | 1449 | 9 | BC011460 | U58515 Human chit1 |
| 27 | 380.6 | 22.6 | 1500 | 9 | HSU58515 | AR042835 Sequence |
| 28 | 360.6 | 21.5 | 1526 | 6 | AR042835 | AF290004 Homo sapi |
| 29 | 338.2 | 20.1 | 1525 | 9 | AF290004 | AF154571 Mus muscu |
| 30 | 311.6 | 18.5 | 1526 | 10 | AF154571 | AF290003 Mus muscu |
| 31 | 311.6 | 18.5 | 1530 | 10 | AF290003 | BC011134 Mus muscu |
| 32 | 311.6 | 18.5 | 1538 | 10 | BC011134 | U15048 Mesocricetu |
| 33 | 266.4 | 15.8 | 2353 | 10 | MAU15048 | D32218 Hamster mRN |
| 34 | 266.4 | 15.8 | 2387 | 10 | HAMOGP | E09047 CDNA encodi |
| 35 | 264.8 | 15.8 | 2366 | 22 | E09047 | U09550 Human ovidu |
| 36 | 258.4 | 15.4 | 2198 | 9 | HSU09550 | U87259 Macaca mula |
| 37 | 253.6 | 15.1 | 2237 | 9 | MMU87259 | M59903 Papio hamad |
| 38 | 250.4 | 14.9 | 2228 | 9 | BABEDOSG | D32137 Mouse mRNa |
| 39 | 242.4 | 14.4 | 2525 | 10 | MUSOGP | E09046 CDNA encodi |
| 40 | 240.2 | 14.3 | 2504 | 22 | E09046 | M94584 Mus musculu |
| 41 | 235.4 | 14.0 | 1525 | 10 | M94584 | D87757 Mus musculu |
| 42 | 233.8 | 13.9 | 1506 | 10 | D87757 | AB025008 Homo sapi |
| 43 | 233.4 | 13.9 | 1354 | 9 | AB025008 | U16719 Ovis aries |
| 44 | 229.4 | 13.6 | 2034 | 4 | OAUI6719 | U17988 Ovis aries |
| 45 | 223.8 | 13.3 | 1596 | 4 | OAUI7988 | |

ALIGNMENTS

| | |
|------------|----------------------------------------------------------------------------------------------------------------------------------------------------|
| RESULT | 1 |
| HUMHA3G | HUMHA3G |
| LOCUS | Human glycoprotein mRNA, complete cds. |
| DEFINITION | M80927 |
| ACCESSION | VERSION |
| KEYWORDS | glycoprotein. |
| SOURCE | Homo sapiens cDNA to mRNA. |
| ORGANISM | Homo sapiens |
| REFERENCE | 1 (bases 1 to 1741) |
| AUTHORS | Hakala,B.E., White,C. and Recklies,A.D. |
| TITLE | Human cartilage gp-39, a major secretory product of articular chondrocytes and synovial cells, is a mammalian member of a chitinase protein family |
| JOURNAL | J. Biol. Chem. 268 (34), 25803-25810 (1993) |
| MEDLINE | 94064658 |
| FEATURES | Location/Qualifiers |
| source | 1. .1741 |

| | | | | | |
|--------------|--------------------------------------------------------|------|------------------------------------------------------------------|----------------------------------------------------------|------|
| 5'UTR | /organism="Homo sapiens" | 601 | cactgtctg | ggggaaggtcaccattgacagcagctatgacattgcagagatatcccaac | 660 |
| sig_peptide | /db_xref="taxon:9606" | 601 | CAC | TGCTCGGGGAAGGTCACCATTTACAGCAGCTATGACATTTGCAAGATATCCCAAC | 660 |
| CDS | /cell_type="chondrocyte" | 661 | acctggatttcattagc | atcatgacacgattttcatggcgctgagctgggaccacag | 720 |
| | /tissue_type="cartilage" | 661 | ACCTGGATTTCATTAGCATCATGACCTACGATTTTCATGGAGCTGCGGTGGGACCCACG | 720 | |
| | 1. .71 | 721 | gcatacacagtc | ccccccagcgaggtcagagatgcaagtctctgacagattcagcaaca | 780 |
| | 72. .134 | 721 | GCCATCACAGTCCCTCTTCCGAGGTGAGGATGCAAGTCTCTGACAGATTTCAGCAACA | 780 | |
| | 72. .1223 | 781 | ctgactatgctg | ggggtacatgttgaggctgggggctctgcagtaagtgcgtggtgag | 840 |
| | /note="articular 39kDa glycoprotein" | 781 | CTGACTATGCTGGGGTACATGTTGAGGCTGGGGGCTCTGCGCAGTAAGCTGTGTATGG | 840 | |
| | /codon_start=1 | 841 | gcatacccaaccttc | gggagagagcttcaactctggtttcttctgaactgggttccagcgc | 900 |
| | /evidence="experimental" | 841 | GCATCCCAACCTTCGCGAGGAGCTTCACTCTGGCTTCTTCTGAGACATGGTGTGGAGCCC | 900 | |
| | /product="glycoprotein" | 901 | caatctcagac | ccgggaattccagcggttcacaaaggagcaggacccttgcctact | 960 |
| | /protein_id="AAAI6074.1" | 901 | CAATCTCAGGACCGGGAAATCCAGCGGTTCCACAAAGGAGGAGGACCTTGCCTACT | 960 | |
| | /db_xref="GI:348912" | 961 | atgagatcgtg | actctccgcggagccacagtcacatagaaacctgcgcagcaggtcc | 1020 |
| | /translators="MGVKA | 961 | ATGAGATCTGTGACTTCTCCGCGGAGCACAGTCCATAGAACCTTCGCGCAGCAGGTCC | 1020 | |
| | SOTGFVVLVLQCCSKYLVCYITWSQYREGDSCPPD | 1021 | cctatgcaccaag | gggcaaccagctgggttagatacagcagcagcaggaagcgtcaaaaagca | 1080 |
| | ALDRFLCTIIYSFANISDNHDIWEWNDVTLKRNPNLKHFTLLISVGGWNFG | 1021 | CCTATGCCACCAAGGGCAACAGTGGGTAGGATACGACACCCAGGAAAGCGTCAAAAGCA | 1080 | |
| | SQRFKSTANTQSRFFIKSVPELTHGFDGLDLAWLPGRDRDKHFTLLIKEMKAE | 1081 | aggtcagctac | ctaagatagcagctggcaggcgccatggtatgggccccctggacctgg | 1140 |
| | FIKEAQPKKOLLISALSGKVTIDSSYDIAKISOHLDPISIMTYDFHGANGTTGH | 1081 | AGGTGCGATGACTGAAGGATAGGACAGCTGGCAGGCGCCATGTTATGGGCCCTGGACCTGG | 1140 | |
| | HSPFLFGQEDASDFRSNTDYAGVLMRLGAPASKLVMGTPTEGRSTLASSTGVA | 1141 | atgactctcag | gggtcctctgctgggcccaggatctgcgcttccctctcaccatgccaatgccatca | 1200 |
| | PISGPGTGPFTKEATLGYEIDFLRGATVHRTLGQQPYATKGNQWGYDDQESV | 1141 | ATGACTTCCAGGGCTCTTCTCGGCCAGAGATCTGCGTCTCCCTCTCAACAAATGCCATCA | 1200 | |
| | KSKVQYLKDLQRLAGAMWALDLDDFGSGCGDLRFLPLTNAIKDALAAT" | 1201 | aggatgcactc | gtaagatagcagctgtcttcacacagcagcggggcccagagatgc | 1260 |
| | 135. .1220 | 1201 | AGGATGCATCTCGATCGAAGTACAGCTTCTGTTCTGGACACAGCAGGGGCCAAGGATGC | 1260 | |
| mat_peptide | /product="glycoprotein" | 1261 | cccgccccctg | ctg-----gctggcgggagcctgatacaccctgcctctgctgagtcgccag | 1314 |
| 3'UTR | 1224. .1741 | 1261 | CCCGTCCCCCTCTGGCTCCAGTCCAGCTGGCGCGGAGCTGATCACCTGCCTGCTGAGTCCCCAG | 1320 | |
| polyA_signal | 1715. .1720 | 1315 | gctgagcctcag | ctcctcccttggggcctatgcaagaggtccacaacacagatttga | 1374 |
| polyA_site | 1741 | 1321 | GCTGAGCCCTCAGTCTCCCTCCCTTGGGGCTATGCAAGGTCCACAAACACACAGATTTGA | 1380 | |
| BASE COUNT | 416 a 510 c 447 g 368 t | 1375 | gctcagcctcgtgggcagagag----- | 1397 | |
| ORIGIN | | 1381 | GCTCAGCCCTTGGTGGGACAGAGGTAGGGATGGGGCTGTGGGGATAGTGAGGCATCGCAA | 1440 | |
| | | 1398 | -----gtacacacttgggtgattgattatgggaaatgtttacagatcccc | 1440 | |
| | | 1441 | TGTAAGACTCGGGATTAGTACACATTTGTTGATGATTAATGGAATGTTTACAGATCCCC | 1500 | |
| | | 1441 | aagcctggcaag | ggaatttcttcaactccccctgccccctgccccctcttatacaggacac | 1500 |
| | | 1501 | AAGCCTGGCAAGGAAATTTCTTCAACTCCCTGGCCCCCTAGCCCCCTTATTAAGAGGACAC | 1560 | |
| | | 1501 | cattttggcag | ctctatcaccaggagccaacatccttacaagacacagtgacca tact | 1560 |
| | | 1561 | CATTTTGGCAAGCTCTATACACAGGAGCCAAACATCTCTACAAGACACAGTGACCATACT | 1620 | |
| | | 1561 | aattatacccc | cttcgaagcagcttgaaaccttcaacttaggaaacgtaactgtgtccccct | 1620 |
| | | 1621 | AATTTATACCCCTGCAAGCCAGCTTGAACCTTCACTTAGGAAACGTAATGCTGCCCT | 1680 | |
| | | 1621 | atcctaacttcccc | cttctaataaccagctgctcaataaagatacaagagtttaacagtgct | 1680 |

```
|||||
Db 1681 ATCCCTACTTCCCTTCCCTAAATCCACAGCTGCTCAATAAGTACAGAGTTTAACAGTGT 1740
Qy 1681 g 1681
Db 1741 G 1741

RESULT 2
LOCUS BC008568 1801 bp mRNA PRI 19-JUL-2001
DEFINITION Homo sapiens, Similar to chitinase 3-like 1 (cartilage glycoprotein-39), clone MGC:17093 IMAGE:4212748, mRNA, complete cds.
ACCESSION BC008568
VERSION BC008568.1 GI:14919432
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1801)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: villalona@bcm.tmc.edu.
Villalona, D.K., Luna, R.A., Hale, S.M., Hulyk, S., Lu, X., Garcia, A.M., Holloway, M., Telford, B., Hodgson, A., Bouck, J., Yu, W., Muzny, D.M., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 12 Row: m Column: 14
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4557017.
FEATURES
source
1. 1801
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:17093 IMAGE:4212748"
/tissue_type="Brain, glioblastoma with EGFR amplification"
/clone_lib="NCI CGAP_Brn64"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
118..1269
/codon_start=1
/product="similar to chitinase 3-like 1 (cartilage glycoprotein-39)"
/protein_id="AAH08568.1"
/db_xref="GI:14919433"
/translation="MGVKSQGTGFVVLVLQCCSAYKLVCVYTSNSQYRGDGSFPD
ALDRFLCTHIYSFANISNDHIDTWENDVTLVGLNLTNLRNPNTLTLISVGNKFE
SQRFSKASNTQRTTKSVPPFLRTHFGDGLAWLYPGRRDKQHTTLIKEMKAE
FIKEAQPKKOLLISAAISAGKVTIDSYDIKISQHLDFISIMYDFHGAWRCTGH
HSPLEFRQEDASPRFNTDYAVGYMLRGLGAPASKLVMGIPTEGRSFTLASSENGVA
PISGPGIPGRTEKAGTLAYEICDFLRGATVHRLIQQVPYATKGNQWGYDDQESV
KSKVQYLRKQLQAGAWWALDLDFQGSFCQDURFLPTNAIKDALAAT"
BASE COUNT 450 a 518 c 462 g 371 t
ORIGIN
```

```
Query Match 91.8%; Score 1543.2; DB 9; Length 1801;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1648; Conservative 0; Mismatches 28; Indels 58; Gaps 3;

Qy 1 ctaggtagctggcaccaggagccgtgggcaagggaagggccacacccctgcctctctg 60
Db 47 CTAGGTAGCTGGCACCAGGAGCCGTGGGCAAGGGAAGAGGCACACACCTGCCCTCTG 106

Qy 61 ctgcagccagaatgggtgtgaagcgctctcaaacagggtttgtgtctgtgtgtctcc 120
Db 107 CTGCAGCCAGAAATGGGTGTGAAGCGCTCTCAACACAGGCTTTGTGTGCTCTGCTGCC 166

Qy 121 agtgcgtctgcatacaaaactggctgtactactacacagctggcccgaccgggaag 180
Db 167 AGTGCCTCTCTGCATACAAACTGGGTCTGTACTACACAGCTGGTCCAGTACCCGGGAAG 226

Qy 181 gcgatgggagctgctccagatgcccttgaccgttctcctgtgtaccacatcatctaca 240
Db 227 GCGATGGGAGCTGCTTCCAGATGCCCTTGACCGCTTCTCTGTATACCCACATCATCTACA 286

Qy 241 gctttgccaatataagcaaacgatcacatcgacacccctggagtggaatgatgtgacct 300
Db 287 GCTTTGCCAATATAAGCAACGATCATCATCGACACCTGGGAGTGAATGATGTGACGCTCT 346

Qy 301 acggcatctcaacacactcaacaacacgaaccccaacccctgaagactctctgtgtgc 360
Db 347 ACGGCATGCTCAACACACTCAAGAACAGGAACCCCAACCTGAAGACTCTCTTTGTCTGTGC 406

Qy 361 gagatggaaactttgggtctcaaaagattttccaagatagctcccaaccccgagctgcc 420
Db 407 GAGGATGGAACCTTTGGGTCTCAAGATTTTCCAGATAGCTTCCAAACACCCAGAGTCCGC 466

Qy 421 ggaacttcaaatcagtaacgcgcattcttgcgcacccatggcttggatggcgctgacc 480
Db 467 GGACTTTTCATCAAGTCACTAGTCAGCGCCATTCTCGCGCACCATGCTTGTATGGGCTGGACC 526

Qy 481 ttgcctgctctaccctgagcgagagacaaacacattttaccacccctaatcaaggaaa 540
Db 527 TTGCCTGGCTCTACCTCGAGCGGAGAGACAAACAGCATTTTACCACCTTATCAAGAGAAA 586

Qy 541 tgaaggccgaatttataaaggaaagcccagcagggaagaaagcagctcctgtcagcgag 600
Db 587 TGAAGCCGGAATTTATAAGGAAGCCAGCCAGGAGGAAAAAGCAGCTCTCTGCTCAGCGCAG 646

Qy 601 cactgtctcgggggaaggtcaccattgcacagcagctatgacattgccaaagatatcccaac 660
Db 647 CACTGTCTCGCGGGAAGGTCAACATTGACAGCAGCTATGACATTGCCAAGATATCCCAAC 706

Qy 661 acctggatttcattagcattacgtacgtacgtatttcattggcgctggcgctgggaccag 720
Db 707 ACCTGGATTTTCATTAGCATCATGACCTACGATTTTCATGGAGCTGGGCTGGGACACAG 766

Qy 721 gccatcacagtccctcaggcgaggtcaggaggtatgcgaagtctgcacagattcagcaaca 780
Db 767 GCCATCACAGTCCCTCTGTCGAGGTTCAGGAGGATGCAAGTCTCTGACAGATTTCAGCAACA 826

Qy 781 ctgactatcgtgtgggtgtacattgttgagctgggggctcctgcacagtaagctgggtgag 840
Db 827 CTGACTATCTGTGGGGTACATGTTTGAAGCTGGGGGCTCTCTGCCAGTAGTAAAGCTGGGTGG 886

Qy 841 gcataccacaccttcgggagaggttcactctgtgtcttcttgagactggtgtgtccagcgc 900
Db 887 GCATCCCCACCTTCGGGAGGAGCTTTCATCTGTGGCTTCTTCTGAGACTGGTGTGTGAGGCC 946

Qy 901 caatctcaggacgggaattccagcgccggttcaccaaggagcagggacccttgcctact 960
Db 947 CAATCTCAGGACCGGAATTCAGSCCGGTTTCAACCAAGSAGSAGGACCCTTGCCTACT 1006

Qy 961 atgagatctgtgactctctccgcggagccacagtcacatagaaacctcggccagcggtcc 1020
Db 1007 ATGAGATCTGTGACTTCTCCGCGGAGGCACAGTCCATAGAAATCCTCGGCCAGCAGGTCC 1066
```


| | | | |
|-----------------------|------------------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|-----------------------------|
| QY | 852 | ttcggagagagcttcactctgtcttctctagactggtgttccagcgcccaatctcagg | 911 |
| Db | 781 | ttcggagagagcttcactctgtcttctctagactggtgttccagcgcccaatctcagg | 840 |
| QY | 912 | ccgggaattccaggcgggttaccacagagcgaggaccccttgctactactagagatctgt | 971 |
| Db | 841 | ccgggaattccaggcgggttaccacagagcgaggaccccttgctactactagagatctgt | 900 |
| QY | 972 | gacttctctcggcgagcacacagtcataagaacctctgcgcagcaggttcccttatgccacc | 1031 |
| Db | 901 | gacttctctcggcgagcacacagtcataagaacctctgcgcagcaggttcccttatgccacc | 960 |
| QY | 1032 | aagggaaccagtggttaggtacacagcagcaagcgtaacgctcaaaagcaggtcaggtac | 1091 |
| Db | 961 | aagggaaccagtggttaggtacacagcagcaagcgtaacgctcaaaagcaggtcaggtac | 1020 |
| QY | 1092 | ctgaaggatagcagctggcggcgccatggtatggcgctggagcctggagctggatctccag | 1151 |
| Db | 1021 | ctgaaggatagcagctggcggcgccatggtatggcgctggagcctggagctggatctccag | 1080 |
| QY | 1152 | ggctctctcggcgagatctgccttccctctcacaatgccaatgccaatgccaatgccaat | 1211 |
| Db | 1081 | ggctctctcggcgagatctgccttccctctcacaatgccaatgccaatgccaatgccaat | 1140 |
| QY | 1212 | gctgcaacg | 1220 |
| Db | 1141 | gctgcaacg | 1149 |
| RESULT | 4 | | |
| LOCUS | E01501 | 1149 bp | RNA |
| DEFINITION | cDNA encoding new polypeptide. | | |
| ACCESSION | E01501 | | PAT |
| VERSION | E01501.1 | GI:2169757 | 29-SEP-1997 |
| KEYWORDS | JP 1988023898-A/2. | | |
| SOURCE | Homo sapiens. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. | | |
| AUTHORS | Furuya,T., Notake,M., Nomura,H., Yamayoshi,M. and Yamada,M. | | |
| TITLE | NOVEL POLYPEPTIDE, DNA CODING SAID POLYPEPTIDE AND PRODUCTION | | |
| JOURNAL | Patent: JP 1988023898-A 2 01-FEB-1988; | | |
| COMMENT | DAINIPPON PHARMACEUT CO LTD | | |
| | OS | human | |
| | PN | JP 1988023898-A/2 | |
| | PD | 01-FEB-1988 | |
| | PF | 16-JUL-1986 | JP 1986167518 |
| | PI | FURUYA TAJIJI, NOTAKE MITSUE, NOMURA HIDEKI, YAMAYOSHI MICHIKO, | |
| | PI | YAMADA MASAAKI | |
| | PC | C07K13/00,C12N15/00,C12P21/02//A61K37/02,(C12P21/02,C12R1:19); | |
| | CC | strandedness: Double; | |
| | CC | topology: Linear; | |
| | CC | hypothetical: No; | |
| | CC | anti-sense: No; | |
| | CC | *source: cell_line=HL-60 cell; | |
| | CC | Feature is identified by experimental; | |
| | FH | Key | Location/Qualifiers |
| | FH | | |
| | FT | CDS | 1..1149 |
| | FT | | /product-'new polypeptide'. |
| FEATURES | source | | |
| | | Location/Qualifiers | |
| | | 1..1149 | |
| | | /organism='Homo sapiens' | |
| | | /db_xref='taxon:9606' | |
| BASE COUNT | 271 a | 333 c | 303 g 242 t |
| ORIGIN | | | |
| Query Match | 66.4% | Score 1115.4; | DB 6; Length 1149; |
| Best Local Similarity | 98.2% | Pred. No. 7e-279; | |
| Matches 1128; | Conservative | 0; Mismatches | 21; Indels 0; Gaps 0; |

| | | | |
|----|------|--------------------------------------------------------------------|------|
| QY | 72 | atgggtggaagcgctctcaaacaggcttttggtggtcctggtgctgctccagtgctgctct | 131 |
| Db | 1 | ATGGGTGTGAAGGCGCTCTCAAAACAGGCTTTGTGGTCTGCTGTCCAGTGTCTCT | 60 |
| QY | 132 | gcatacaaacctggtctgtctactacacccagctggttcccagtagcgggaagcgatgggagc | 191 |
| Db | 61 | GCATACAACATGGTCTGCTACTACACAGCTGTGCCAGTACCGGNAAGGCATGGAGC | 120 |
| QY | 192 | tgcttcccagatggccttgaccgcttccctgtgttaccacaatcatctacagcttttccaat | 251 |
| Db | 121 | TGCTTCCAGATGCCCTTGACCGCTTCTGTACCCACATCTGTACCCACATCATACAGCTTTGCCAAT | 180 |
| QY | 252 | ataaacaacgatacactcgacacacctggagtggaatgagatgtgacactctacagatgctc | 311 |
| Db | 181 | ATAACAACGATACATCGACACCTGGGAGTGGATGATGTGACCTCTTACGGCATGCTC | 240 |
| QY | 312 | aacacactcaacaacacgaaaccccccaacctgaagactctcttctgtcggagagatggaaac | 371 |
| Db | 241 | NACACACTCAAGAACACAGGAACCCCAACCTCAAGACTCTCTTCTGTCTGGAGATGGAAC | 300 |
| QY | 372 | tttgggtctcaaaagattttccaaagatagcctccaaacacccagagtcgcggagcttttcac | 431 |
| Db | 301 | TTTGGGTCTCAAAAGATTTTCCAAGATAGCTTCCAACACCCAGAGTCGCCGAGCTTTTCATC | 360 |
| QY | 432 | aagtcagtagcgcgaatttctgcgcacctggtctttagtggtggtgacaccttgcctggtc | 491 |
| Db | 361 | AAGTCAGTACCCGCCATTTCTGCGCACCCATGGCTTTGTATGGGCTGGACCTTGCTGGGCTC | 420 |
| QY | 492 | tacctggagcgagagacaacacacctttttaccacctaatcaaggaatgaagcgccgaa | 551 |
| Db | 421 | TACCTGGACGGGAGACAAACAGCATTTTACCACCTTAATCAAGGAATGAAGGCCGAA | 480 |
| QY | 552 | tttataaaggaaagccagccagggaaaaagcagctcctgtcagcgcagcactgtctgcy | 611 |
| Db | 481 | TTTATAAAGGAAGCCAGCCAGGGAAGAAAGCAGCTCTCTCAGCGCAGCACTGTCTGG | 540 |
| QY | 612 | gggaaggtcaccattgacagcagctatgacattgccaagatattcccaacacctggatttc | 671 |
| Db | 541 | GGGAAGGTCAACATTTGACAGCAGCTATGACATTTGCCAAGATATCCCAACACCTGGATTTC | 600 |
| QY | 672 | attagcatcatgacatagctatttcatggcgctgctgctcagcgcagcactgtctgcy | 731 |
| Db | 601 | ATTAGCATCATGACCTACGATTTTTCATGGAGCTGCGTGGGACCCACACAGGCCATCACAT | 660 |
| QY | 732 | cccctcagcgaggttcaggaggtgcaagtcctgcagattcctgcagcaacactgactatgct | 791 |
| Db | 661 | CCCCTGTTCCGAGGTCAGGAGGATGCAAGTCTCTGACAGATTGACAAACACTGACTATGCT | 720 |
| QY | 792 | gtgggtacatgttgaggtggggctcctgcagtaagctggtgagtggtggtggtggtggtggt | 851 |
| Db | 721 | GTGGGTACATGTTGAGGTGGGGCTCCTGCCAGTAAAGCTGCTATGGGATGCCCCACC | 780 |
| QY | 852 | ttcgggagagcttcaacttgcttctcttgagactggtgttccagcgccaaatctcagg | 911 |
| Db | 781 | TTGCGGAGGAGCTTCACCTCTGCTTCTGTGAGACTGGTGTGGAGCCCCCAATCTCAGGA | 840 |
| QY | 912 | ccgggaattccagcggtttcaccacagggagggagcccttgcctactatgagatctgt | 971 |
| Db | 841 | CCGGGAATTCACAGGCCGGTTTACCAGGAGGCGAGGCCCTTGCTTACTATGAGATCTGT | 900 |
| QY | 972 | gacttctcgcggagcacacagtcataagaacctctgcgcagcaggttcccttatgccacc | 1031 |
| Db | 901 | GACTTCTCCGCGAGCCACAGTCCATAGAAATCTTCCGCGCAGAGGTCCCTTATGCCACC | 960 |
| QY | 1032 | aagggaaccagtggttaggtacacagcagcaagcgtaacgctcaaaagcaggtcaggtac | 1091 |
| Db | 961 | AAGGGCAACAGTGGGTAGGATACGACACCCAGGAAGAGGCTCAAAAGCAAGGTGCGAGTAC | 1020 |
| QY | 1092 | ctgaaggatagcagctggcggcgccatggtatggcgctggagcctggagctggatctccag | 1151 |
| Db | 1021 | CTGAAGGACAGCAGCTGGCGGGGCCCATGGTATGGGCCCTGGACCTGGATGACTTCCAG | 1080 |

| | |
|------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| RESULT | 9 |
| BC005611 | |
| LOCUS | 1564 bp mRNA 12-JUL-2001 |
| DEFINITION | Mus musculus, Similar to chitinase 3-like 1 (cartilage glycoprotein-39), clone MGC:7884 IMAGE:3582304, mRNA, complete cds. |
| ACCESSION | BC005611 |
| VERSION | BC005611.1 GI:13542819 |
| KEYWORDS | MGC. |
| SOURCE | house mouse. |
| ORGANISM | Mus musculus |
| | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. |
| REFERENCE | 1. (bases 1 to 1664) |
| AUTHORS | Strausberg, R. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (02-APR-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590. |

| | | | | | |
|--------------|------|---------------------------------------------------------------|------|--------|--|
| Db | 1385 | GTGTGAGAGTCACAGTGTGAGCAGATACACAAACCCCTGTAAAGGAATGCAAAATCTCAGA | 1444 | ORIGIN | |
| Qy | 1466 | ctccctgcccc--tagccctcttataaaagacacacattttggcaagctctatcacca | 1523 | | |
| Db | 1445 | CTCTAACCTCCCTTTACCCAGAGCTGACCAAGAGACACCACTTGG-----ATCA | 1492 | | |
| Qy | 1524 | agagccaaacatctacaagacacagtgacataactaattatccccctgcaag-cca | 1582 | | |
| Db | 1493 | AGTAGGCAAAATCTTTACAGGATTGAGGACCATACTAATTATACCTCTGCAAGCCCA | 1552 | | |
| Qy | 1583 | gcttgaacaccttaccttagaagcgtatgctcccttacttacttcccccttctaatt | 1642 | | |
| Db | 1553 | ACTGTGATCTCTCCCTTAGGAACCTTAATCGTCCCACTTCCCTTTC-----CCTAAT | 1604 | | |
| Qy | 1643 | ccacagctgtcctaataaagataagagtttaaca | 1676 | | |
| Db | 1605 | CCACAGCTGTTCATTAAGGCCGAGAACCTTAAAA | 1638 | | |
| RESULT 10 | | | | | |
| MMBRP39 | | | | | |
| LOCUS | | | | | |
| DEFINITION | | | | | |
| ACCESSION | | | | | |
| VERSION | | | | | |
| KEYWORDS | | | | | |
| SOURCE | | | | | |
| ORGANISM | | | | | |
| REFERENCE | | | | | |
| AUTHORS | | | | | |
| TITLE | | | | | |
| JOURNAL | | | | | |
| MEDLINE | | | | | |
| REFERENCE | | | | | |
| AUTHORS | | | | | |
| TITLE | | | | | |
| JOURNAL | | | | | |
| FEATURES | | | | | |
| source | | | | | |
| misc_feature | | | | | |
| gene | | | | | |
| CDS | | | | | |
| BASE COUNT | | | | | |

Query Match 44.7%; Score 751; DB 10; Length 1616;
Best Local Similarity 71.8%; Pred. No. 3.2e-184; Indels 53; Gaps 9;
Matches 1128; Conservative 0; Mismatches 390;


```
Db 990 GTACATCGACTCTCCAAACGAGAAGGTTCCCTTCCTGCTACCAAGGCCAACCCAGTGGGTGGG 1049
QY 1053 taccagaccaggaagagctcaaaagcgaaggctgacagtaactgaagatagcagctggca 1112
Db 1050 TATGACCAAGAGAGTGTCAAAACAAAGGTTGGGTTCTCTGAAGGAGAAAGCTGGCA 1109
QY 1113 ggcgcctatggtatgggcccctgacccctgagcttccagctccagggctccctctcgcgcagat 1172
Db 1110 GGAGCCATGGTGGGCACCTGGATTTGGATGATTTCAGGGCACCTGTACGCCGAGGAA 1169
QY 1173 ctgcgcttcctctcaccatgccatcaagatgacactgcctgcctgcaacgtagccctgtt 1232
Db 1170 TT---CTTCCCGCTACCAACGCCATCAAGGATGCTGGCT-----TAGCTCCCTT 1220
QY 1233 ctgcacacagcagcggggcccaaggatgcgccgtccccgtctggctggcggcggagcctgat 1292
Db 1221 TCCCATATGGTA-----CCCCCACTCTCTGGCCAGGAGTTTAAT 1259
QY 1293 cactgcctctgtag-tcccaggtgagctcag--tctccctcccttgggcccctatgc 1349
Db 1260 CTTTCGATGTTAAGTTCCCAACTGAGCCTCAGTTTCTCTTCCCTTGGCACCTGTGT 1319
QY 1350 agaggtccacaacacacaga---ttgagctcagccctggtggcagaggtacacact 1406
Db 1320 AAGGGCCACAGCAGCTCAGCTATGGAACAGGGAACCTAGGGTAGGAGATGGTGGG 1379
QY 1407 tgtgatgattaatgaaatgtttacagatccccagcctggaagggaaattcttcaac 1466
Db 1380 TTGTGAGTCAAGTGTGAGCAGATACAAACCCCTGTTAAGGAATTCGAATTTCTCAGAC 1439
QY 1467 tc---cctgcctcctgacctccttataaaggacacacatttttgcaagctctatcacca 1524
Db 1440 TCTAACCTCCCTTACCAGCCTGACCAAGGACACCCTGG-----ATCAA 1487
QY 1525 gaggccaaacactcctcaagacacagtgaccataactaattatccccctgcgaagccagc 1584
Db 1488 GTAGGCAATATCTTACAGGATTGAGGGACCATACTAATATATACCTCTCTGCAAGGCCAA 1547
QY 1585 ttgaacacctc 1595
Db 1548 CTTGAATCTTC 1558

RESULT 11
AF011373
LOCUS AF011373 997 bp mRNA MAM 31-JUL-1997
DEFINITION Bos taurus chitinase-like protein 1 (CLP-1) mRNA, partial cds.
ACCESSION AF011373
VERSION AF011373.1 GI:2286218
KEYWORDS
SOURCE COW.
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;
Bovidae; Bovinae; Bos.
REFERENCE 1 (bases 1 to 997)
Recklies,A.D. and White,C.
EXPRESSION OF CHITINASE-LIKE PROTEIN 1 (CLP-1) IN BOVINE
CHONDROCYTES
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 997)
Recklies,A.D. and White,C.
AUTHORS
TITLE Direct Submission
JOURNAL Submitted (27-JUN-1997) Joint Diseases Lab, Shriners Hospital, 1529
Cedar Ave, Montreal, QUE H3G 1A6, Canada
FEATURES
source
1..997
/organism="Bos taurus"
/db_xref="taxon:9913"
/cell_type="articular chondrocytes"
<1..>997
/gene="CLP-1"
<1..>997
gene
CDS
```

```
/gene="CLP-1"
/codon_start=1
/product="chitinase-like protein 1"
/protein_id="AAB64304.1"
/db_xref="GI:2286218"
/translation="VCYVTSWQYREGDSCFPDAIDPFLCHTVIYSFANISNEIDT
LWWDVTLVDLNTLTKNRPNKLTLLSVGMWNFQSKIASKTSRRTFKSVPPF
LTHGFDGLDLAWLPGRWKRHLTTLVKEMAEFVRGAQAGTEQLLLSAAVPAGRIA
IDRGYDIAQISRLDFISLLTYDFHCAWROTVGHHSPLERGOEDASSDRFSNADYAVS
YMLRGAPANKLVMGIPITFGRSYTLASSKTVDGAPISGPGIPQGFQKRGILAYEIC
DFLHGATTHFRDQVYATKGNQWVAYDDQESVKNKARYLKNROLAGAMVWALDL"
BASE COUNT 248 a 281 c 267 g 201 t
ORIGIN
Query Match 44.6%; Score 749; DB 4; Length 997;
Best Local Similarity 84.5%; Pred. No. 1.le-183;
Matches 842; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
QY 144 gtctgtactacacagctggtccagtagccggagagcgatggagctgttcccagat 203
Db 1 GTCTGTTATTACACCAGCTGGTCCCAGTACCGGAGGGGTATGGAGCTGCTTCCCAGAC 60
QY 204 gccttgaccgtctctgtgtaccacacatctctacagctttgccaataaagcaacgat 263
Db 61 GCCATCGACCCCTTCCTGTGCACCCATGTCTATACAGCTTGGCCACATTAAGCAACAT 120
QY 264 cacatgcacacctggagtggaatgatgtgacgctctacgcatgctcaacacactcaac 323
Db 121 GAGATCGACACCTGGGAGTGGAAATGAGCTGACGCTCTATGACACACTGAACACACTCAAG 180
QY 324 aacacgaaccccaacctgaagactctctgtctgctgagagatggaacttgggtctcaa 383
Db 181 AACAGGAACCCCAACCTTGAAGACCCCTCTATCTGTGTGGAGATGGAACCTTCGGTTCCAA 240
QY 384 agattttccaagatagcctcccaacacacagagctgcgcggacttttcaaaagtcagtcaccg 443
Db 241 AGATTTTCCAAGATAGCTTCCAAAGACCCAGAGTCGCGAGGACTTTTCATCAAGTCGGTGCCA 300
QY 444 ccaattctgcgcacccatggcttggcgctgagaccttgcctggctctaccctggacgg 503
Db 301 CCATTTCTGCGGACCCATGGCTTTGTATGGACTGGACCTAGCATGGCTTCTACCCCGGGTGG 360
QY 504 agagacaacacacattttaccacccctaatacaagaaatgaagccgaattataaaggaa 563
Db 361 AGAGACAAGCGGATCTCACCACTCTGGTCAAGAAATGAAGGCTGAGTTGTAAAGGGGA 420
QY 564 gccacgcagggaaaaagcagctcctgctcagcgacgactgtctgcgggaaggtcacc 623
Db 421 GCCCAAGCAGGCACAGACAGCTTCTGCTCAGTCGACGACTACCTGCAGGAGATTTGCT 480
QY 624 attgacagcagctatgacattgccaagatatacccaacacacctgatttcattagcagcatg 683
Db 481 ATTGACAGAGGCTATGACATCGCCAGATATCCGACACCTGGGACTTCATCAGCAGCTTTTG 540
QY 684 acctagattttcatggcgcttggcgctgggacacacagggccatcacagtcacctcaggcga 743
Db 541 ACCTATGACTTTTACGGAGGCTTGGCCGACAGACAGTTCGCGACACACGCCCTCTTTCGA 600
QY 744 ggtcaggagagatgcaagctcctgacagatttcagcaaacactgactatgctgtgggtacatg 803
Db 601 GCCCAGGAAGATGCAAGTTCTGACAGATTTCAGTAACGCTGACTACGCTGTGAGCTACATG 660
QY 804 ttgagctggggctcctgcagtaagctggtgatgggcatccccaccttcgggagagc 863
Db 661 CTGAGGCTGGGGCTCCAGCCAATAAGCTGTTGATGGGTATCCCACTTTTGGGAGGAGC 720
QY 864 ttcaactctggctctctctgagactggtttccagcgccaatctcaggacccaggggaattcca 923
Db 721 TACACTCTGGCCCTCTCCAAAGACAGATGTGGGAGCCCCCATCTCAGGGCCAGGAATTCGA 780
QY 924 ggcsggttccacaagagcgaggaccttgcctactatgactatgactgtgacttctccgc 983
```

Tue Dec 11 08:46:51 2001

us-09-164-862b-4.rge

```

Db 781 GGCAGTTCACCAAGAGAGAAAGGATCTTGCCTATTATGAGATCTGTGACTTCCTCCAC 840
QY 984 ggagccagtcattatgaacacctcgccagcaggtccctctatgcccacaaagggcaccag 1043
Db 841 GGAGCCACCACCCACAGATTCCGTGACCAAGAGTCCCTATGTCACCAAGGCAACACAG 900
QY 1044 tgggtaggatatacaccacaggaagaaagcgtcaaaagcaaggtgagtcagctgaaggtatgg 1103
Db 901 TGGGTGGCGTATGACGACCAAGGAGAGTGTCAAAAACAAAGCAGCGGTACCTGAGAACAGG 960
QY 1104 caactgacagccctggtatgagccctggacctgg 1140
Db 961 CAGTGGCTGGCGCATGTTATGGCCCTGGACCTGG 997

RESULT 12
E01502 966 bp RNA PAT 29-SEP-1997
LOCUS cDNA encoding new polypeptide.
DEFINITION E01502
ACCESSION E01502
VERSION E01502.1 GI:2169758
KEYWORDS Homo sapiens.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 966)
AUTHORS Furuya,T., Notake,M., Nomura,H., Yamayoshi,M. and Yamada,M..
TITLE NOVEL POLYPEPTIDE, DNA CODING SAID POLYPEPTIDE AND PRODUCTION
JOURNAL Patent: JP 198023898-A 3 01-FEB-1988;
DAINIPPON PHARMACEUT CO LTD
COMMENT OS human
PN JP 1988023898-A/3
PD 01-FEB-1988
PF 16-JUL-1986 JP 1986167518
PI FURUYA TAJIJI, NOTAKE MITSUE, NOMURA HIDEKI, YAMAYOSHI MICHIO,
PC C07K13/00,C12N15/00,C12P21/02//A61K37/02,(C12P21/02,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: cell_line=HL-60 cell;
CC feature is identified by experimental;
FH key Location/Qualifiers
FT CDS 1..966
FT /product='new polypeptide'.
FEATURES
source 1..966
Location/Qualifiers
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 238 a 285 c 244 g 199 t
ORIGIN

Query Match 44.3%; Score 744.2; DB 6; Length 966;
Best Local Similarity 82.5%; Pred. No. 1.9e-182;
Matches 948; Conservative 0; Mismatches 18; Indels 183; Gaps 1;

QY 72 atgggtgtagcgctctcaaacagcgtttgtgctcctggtgctgctcagtgctctct 131
Db 1 ATGGGTGTCAGGCGTCTCAACAGCGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTCT 60
QY 132 gcatacaaacctggtctgctactacacagctggtccagtcacccggaagcgatggagc 191
Db 61 GCATACAAACTGCTGCTACTACACAGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTG 120
QY 192 tgcctccagatgacctgacgcttctgtgtacccacatcatctacagctttgccaat 251
Db 121 TGCCTCCAGATGCCCTTGACCGCTTCTCTGTAGCCACATCATCTACAGCTTTGCCAAT 180
QY 252 ataagcaacgatcacatcgacacacctggagtggaatgatgtgacgctctcagcgatgctc 311

```

RESULT 13
AF062038
LOCUS
DEFINITION
ACCESSION

AF062038 1056 bp mRNA ROD 02-APR-1999
Rattus norvegicus glycoprotein-39 precursor, mRNA, partial cds.
AF062038

```

Db 181 ATAAGCAACGATCATCATGACACCTGGAGTGAATGATGTGACCTCTACGGCATGCTC 240
QY 312 aacacactcaacacacacacaccccaacctgaagactctcttctgtctgtgagagatggaaac 371
Db 241 AACACACTCAAGAACAGGAACCCCAACCTGAAGACTCTCTTGTCTGTGGAGGATGGAAC 300
QY 372 tttgggtctcaagatattttccaagatagcctccaacacccacagagtcgcgcgactttcacc 431
Db 301 TTTGGGTCTCAAGATATTTTCCAAAGATAGCTCCAACACCCAGAGTCGCCGGACTTTTCATC 360
QY 432 aagtcagtaacccctattctcgacccctgcttggatggcggtgagccttgcctggctc 491
Db 361 AAGTCAGTACCGCCATTCTTGCACCCCATGGCTTGTATGGGCTGGACCTTGCCTGGCTC 420
QY 492 taccctggacggagagacaaacacacattttacacccctaatcaaggaatgaagggccgaa 551
Db 421 TACCTTGGAGGGGAGACAAACAGCATTTTACACCCCTAATCAAGGAAATGAAGGCCGAA 480
QY 552 ttataaaggaagcccgagccaggggaaagacagctcctgctcagcgacgacactgtctg 611
Db 481 TTTTATAAGGAAGCCAGCCAGGCAAGGAAAAGCAGCTCTGCTCAGCGCAGCAGCTGTCTGG 540
QY 612 ggaaggtcaccattgacacagcagctatgacattgccaagatatcccaacacctggatttc 671
Db 541 GGAAGGTCAACATTGACACAGCAGCTATGACATTGCCAAGATATCCCAACACCTGGATCTC 600
QY 672 attagcatcatgacctacgattttcattgctgctgctgctgctgctgctgctgctgctgct 731
Db 601 ATTAGCATCATGACTACGATTTTTCATGGAGCCCTGGCGTGGGACACAGGCGCATCACAGT 660
QY 732 cccctcagggaggtcagggaggtcaggtcctgacagattcagcagattcagcagacactgact 791
Db 661 CCCCTGTCCGAGGTGAGGAGGATGCAAGTCTCTACAGATTCAGCAACACT----- 711
QY 792 gtgggttacatgttgaggctgggggctcctgcccagtaagctgggtgagtgagtgagtgag 851
Db 712 ----- 711
QY 852 ttcggaggaggttccactgctgctgctgctgctgctgctgctgctgctgctgctgctgctg 911
Db 712 ----- 711
QY 912 ccggaaattccagccggttcaccaagaggcagggacccttgctactatgagatgctgt 971
Db 712 -----ATCTGT 717
QY 972 gacttcctccgcggagccacagtcacagtcacagtcacagtcacagtcacagtcacagtcac 1031
Db 718 GACTTCTCCGCGAGCCACAGTCCATAGATCTCTCGGCCAGCAGGTCCCTTATGCCACC 777
QY 1032 aagggaacacagtggttagatagcagcagcagcagcagcagcagcagcagcagcagcagcag 1091
Db 778 AAGGGCAACCACTGGGTAGGATACGACGACGACGACGACGACGACGACGACGACGACGAC 837
QY 1092 ctgaaggtataggcagctgagcagcagcagcagcagcagcagcagcagcagcagcagcag 1151
Db 838 CTGAGGACAGCAGCAGCTGGCGGCCCATGTTATGGGCCCTGGACCTGGATGACTTCCAG 897
QY 1152 ggcctccttgcggcagagatctgctcctcctcctcctcctcctcctcctcctcctcctcct 1211
Db 898 GGCTCCTTCTGGCGCAGGATCTGCGCTTCCCTCTACCAATGCCATGCCATCAAGGATGCATC 957
QY 1212 gctgcaacg 1220
Db 958 GCTGCAACG 966

```



```

AF062038.1 GI:4558457
.
Norway rat.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.

REFERENCE
AUTHORS      1 (bases 1 to 1056)
TITLE        Wending,U., Boots,A.M.H. and van Eden,W.
JOURNAL      potential autoantigen in arthritis
              unpublished
REFERENCE
AUTHORS      2 (bases 1 to 1056)
TITLE        Wending,U., Boots,A.M.H. and van Eden,W.
JOURNAL      Direct Submission
              Submitted (29-APR-1998) Institute of Infectious Diseases and
              Immunology, University of Utrecht, Faculty of Veterinary Medicine,
              Yalelaan 1, 3508 TD Utrecht, The Netherlands

FEATURES
source       Location/Qualifiers
              1..1056
                /organism="Rattus norvegicus"
                /strain="Lewis"
                /db_xref="taxon:10116"
                /clone="RCP391"
mat_peptide  1..>1056
              /product="glycoprotein-39"
              <1..>1056
              /note="similar to the Homo sapiens product encoded by
              GenBank Accession Number M80927"
              /codon_start=1
              /product="glycoprotein-39 precursor"
              /protein_id="AAD22610.1"
              /db_xref="GI:4558458"
              /translation="YKLVCYTNWQYREGNGSCFPDALDHSICLTHLIYSEANISNKK
              LSTSEWNDVLYGMLNTLRPRXLTLISVGWWSGSRFSRIYVSNKSRKTFVQSV
              APFLTQYDGLDQAWLQYPPKDKQHFHTLILKELAAETKEVQPGTEKLLLSAAVSAG
              KVLTDGSDYDAQIAHLDFINLMYTFDHTWRHTTHGHSPLFRQGDQTPDRFSNDYD
              GYGMRLRGAPTNKLVMGIPFGKSFPTLASSENQVGAPISGSGLPGRYTKKGTLLAY
              EICDFLURGAEVHRLIGQVPPFATKGNQWVGYYDDPESVKNKVYLNKQLAGAMVWVAD
              LDDFRGSFCGHNVHEPLT"

BASE COUNT   275 a   289 c   270 g   222 t
ORIGIN
Query Match      43.8%;   Score 736;   DB 10;   Length 1056;
Best Local Similarity 81.1%;   Pred. No. 2.6e-180;
Matches 856; Conservative 0; Mismatches 200; Indels 0; Gaps 0;

QY  135  tacaactggtctgtctactacacagctggtctccagttaccgggaagcgatggagctgc 194
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   1  TACAACTGGTGCTGCTACTACACCACTGGTGCCAGTACCGGAAGCAATGGGAGCTGC 60

QY  195  ttcccaagatgcccttgagcgccttctctgtgtaccacatctctacagctttgccaata 254
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   61  TTCCACAGCCCTTCGACCAATCTCCTGTGCACCCATCATCTACAGCTTTGCCAACATC 120

QY  255  agaacatgcacatcgacacatcgagtggaatgatgtgacgtctcagcgtactcgtcaac 314
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   121  AGCAACAACAGCTCAGCACATCGGATGGAAAGACGTAAACCTGTATGGCATCTCTGAAT 180

QY  315  aactcaacaacgaaccccaactgaagactctctgtctcgagagatggaacttt 374
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   181  ACTCTAAGACCAGAAACCCAGACTCAAGACACTGCTGCTGTTGGAGGATGGAGCTTT 240

QY  375  gggctctcaagattttccaagatagcctccaacaccagagtcgcggactttcataag 434
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   241  GGCTCAGAAAAGATTTTCCAGGATTTGCTCCAACGCTAAGAGTCGCAAGACTTTTCGTCCAG 300

QY  435  tcagtaccgcattcttcgcaccatggctttgatggcgtagccttgctgctgctac 494
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db   301  TCGGTAGCTCCCTTTCGCGGACCTATGGCTTTGATGGACTGGATCTCGGCTGCTCTAC 360

QY  495  cctggcggagagacaacaccatttttaccacccctaacgaagaaatgaagccgcaattt 554

```


us-09-164-862b-4.rge

Tue Dec 11 08:46:51 2001

• • • • •

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2001, 21:27:35 ; Search time 178.91 seconds
(without alignments)
8055.244 Million cell updates/sec

Title: US-09-164-862B-4

Perfect score: 1681

Sequence: 1 ctggtgtagctggcaccaggga.....tacaagagtttaacagtggtg 1681

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 930621 seqs, 428662619 residues

Total number of hits satisfying chosen parameters: 1861242

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_1101.*
1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|-----------------------------|
| 1 | 1674.6 | 99.6 | 1681 | 16 | AA085245 YKL-40 gene. Homo |
| 2 | 1671.4 | 99.4 | 1681 | 21 | AAZ94901 Human cancer marke |
| 3 | 856 | 50.9 | 1152 | 18 | AA099452 Bovine whey protei |
| 4 | 745.8 | 44.4 | 966 | 9 | AAN81756 Gene encoding poly |
| 5 | 446.2 | 26.5 | 503 | 21 | AAC00233 Human secreted pro |
| 6 | 399.2 | 23.7 | 1713 | 18 | AAT50834 Human chitinase cd |
| 7 | 397.6 | 23.7 | 1643 | 18 | AAT50833 Human chitinase cd |
| 8 | 397.2 | 23.6 | 1636 | 19 | AAV10435 Human chitinase cl |
| 9 | 397.2 | 23.6 | 1636 | 20 | AAZ21847 MO-218 clone of hu |
| 10 | 397.2 | 23.6 | 1636 | 22 | AA03759 Human chitinase cd |
| 11 | 397.2 | 23.6 | 1768 | 18 | AAT89181 Human chitotriosid |

| | | | | | |
|----|-------|------|------|----|-----------------------------|
| 12 | 396 | 23.6 | 1656 | 19 | AAV10436 Human chitinase cl |
| 13 | 396 | 23.6 | 1656 | 20 | AAZ21848 MO-13B clone of hu |
| 14 | 396 | 23.6 | 1656 | 22 | AA03760 Human chitinase cd |
| 15 | 389.2 | 23.2 | 1637 | 18 | AAT89180 Human chitotriosid |
| 16 | 386.2 | 23.0 | 1432 | 19 | AAV21689 DNA encoding a hum |
| 17 | 384.6 | 22.9 | 1433 | 19 | AAV13925 Human cartilage gp |
| 18 | 382.2 | 22.7 | 1496 | 18 | AAT97127 Human cartilage q1 |
| 19 | 362.2 | 21.5 | 1594 | 19 | AAV21687 DNA encoding a hum |
| 20 | 360.6 | 21.5 | 1526 | 19 | AAV13926 Human cartilage gp |
| 21 | 359 | 21.4 | 1594 | 19 | AAV21688 DNA encoding a hum |
| 22 | 339.8 | 20.2 | 1678 | 22 | AAH42025 Disease treatment |
| 23 | 325 | 19.3 | 1368 | 22 | AAH42013 Osteoarthritis tis |
| 24 | 301.8 | 18.0 | 1474 | 22 | AAH23078 Human breast cance |
| 25 | 273.4 | 16.3 | 468 | 22 | AAV17558 Hamster oviduct sp |
| 26 | 264.8 | 15.8 | 2366 | 16 | AAQ90444 Murine oviduct spe |
| 27 | 240.2 | 14.3 | 2504 | 16 | AAQ90443 Disease treatment |
| 28 | 233.8 | 13.9 | 1469 | 22 | AAH42023 Human gene signatu |
| 29 | 229.8 | 13.7 | 251 | 16 | AAT19620 Bovine oviduct spe |
| 30 | 222.4 | 13.2 | 1994 | 16 | AAQ90442 Human secreted pro |
| 31 | 136 | 8.1 | 379 | 21 | AAC00135 Cartilage-associat |
| 32 | 105 | 6.2 | 426 | 21 | AAZ94491 Human EST-derived |
| 33 | 82.2 | 4.9 | 339 | 22 | AAH98755 D. pteronyssius 98 |
| 34 | 77.2 | 4.6 | 1470 | 21 | AAZ38589 D. pteronyssius 98 |
| 35 | 77.2 | 4.6 | 1470 | 21 | AAZ38590 D. pteronyssius 98 |
| 36 | 77.2 | 4.6 | 1527 | 21 | AAZ38587 D. pteronyssius 98 |
| 37 | 77.2 | 4.6 | 1527 | 21 | AAZ38588 D. pteronyssius 98 |
| 38 | 77.2 | 4.6 | 1621 | 21 | AAZ38585 D. pteronyssius 98 |
| 39 | 77.2 | 4.6 | 1621 | 21 | AAZ38586 D. pteronyssius 98 |
| 40 | 76.6 | 4.6 | 1608 | 21 | AAZ38579 D. farinae mite al |
| 41 | 76.6 | 4.6 | 1608 | 21 | AAZ38580 D. farinae mite al |
| 42 | 76.6 | 4.6 | 1665 | 21 | AAZ38577 D. farinae mite al |
| 43 | 76.6 | 4.6 | 1665 | 21 | AAZ38578 D. farinae mite al |
| 44 | 76.6 | 4.6 | 1752 | 21 | AAZ38575 D. farinae mite al |
| 45 | 76.6 | 4.6 | 1752 | 21 | AAZ38576 D. farinae mite al |

ALIGNMENTS

RESULT 1
AA085245
ID AA085245 standard; cDNA; 1681 BP.
XX
AC AA085245;
XX
DT 12-AUG-1995 (first entry)
XX
DE YKL-40 gene.
XX
KW YKL-40; diagnosis; prognosis; therapy; breast cancer; metastasis;
KW marker; joint disease; connective tissue; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 135..1681
FT /*tag= a
XX
PN WO9501995-A.
XX
PD 19-JAN-1995.
XX
PF 08-JUL-1994; 94WO-US07754.
XX
PR 09-JUL-1993; 93US-0089989.
XX
PA (REGC) UNIV CALIFORNIA.
XX
PI Johansen JS, Price PA;
XX
DR WPI; 1995-066866/09.
XX
PT Use of YKL-40 and anti-YKL-40 antibodies - for developing prods.

PT for diagnosis prognosis and therapy of diseases involving
PT connective tissue degradation.

XX Disclosure; Page 66-67; 88pp; English.

XX YKL-40 (40 kDa) was purified from human osteosarcoma MG63 cells.
CC The N-terminal sequence is shown in AAR70745; the full coding region
CC of the YKL-40 gene is given in AAQ85245. Homology of the N-terminal
CC and 2 internal peptides (AAR70746-47) with a bacterial polysaccharide
CC hydrolase suggests that YKL-40 degrades polysaccharide components
CC of connective tissue. YKL-40 is a marker of e.g. metastatic breast
CC cancer and inflammatory or degenerative joint diseases.

XX Sequence 1681 BP; 401 A; 503 C; 421 G; 356 T; 0 other;

Query Match 99.6%; Score 1674.6; DB 16; Length 1681;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1677; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 ctaggtagctggcaccaggccgctgggcaagggaagagccacaccctgctgtctg 60
Db 1 ctaggtagctggcaccaggccgctgggcaagggaagagccacaccctgctgtg 60
Qy 61 ctgcagcagaatgggtgaaggctctcaacagagctttgtgctgtgctgtcc 120
Db 61 ctgcagcagaatgggtgaaggctctcaacagagctttgtgctgtgctgtcc 120
Qy 121 agtgcgtctgcatacaaaactgctgctactacacacagctggctccagtaacgggaag 180
Db 121 agtgcgtctgcatacaaaactgctgctactacacacagctggctccagtaacgggaag 180
Qy 181 gcgatggagctgtctccagatgctctgacagcttctctgtgtacacacatactaca 240
Db 181 gcgatggagctgtctccagatgctctgacagcttctctgtgtacacacatactaca 240
Qy 241 gctttgcaataaagaacacatcacatcgacactggagtggaatgatgtgacgtct 300
Db 241 gctttgcaataaagaacacatcacatcgacactggagtggaatgatgtgacgtct 300
Qy 301 acgggatgctcaacacactcaacacgaaccccaacccctgaagactctctgtgtctg 360
Db 301 acgggatgctcaacacactcaacacgaaccccaacccctgaagactctctgtgtctg 360
Qy 361 gaggatgaactttggctctcaagattttccaagatagctctcaacacccagagtgccc 420
Db 361 gaggatgaactttggctctcaagattttccaagatagctctcaacacccagagtgccc 420
Qy 421 ggaacttcaatgaactcagtcacgcgcatcttctgcacccatggctttgatggcgtagcc 480
Db 421 ggaacttcaatgaactcagtcacgcgcatcttctgcacccatggctttgatggcgtagcc 480
Qy 481 ttgcttggtctaccctggagcagagagacaaacacacattttacacacctaatacaggaaa 540
Db 481 ttgcttggtctaccctggagcagagagacaaacacacattttacacacctaatacaggaaa 540
Qy 541 tgaaggcgaattataaagaagcccgagcaggaagaaagcagctctgctcagcgcag 600
Db 541 tgaaggcgaattataaagaagcccgagcaggaagaaagcagctctgctcagcgcag 600
Qy 601 cactgtctggggagggtccaccattgacagcagctatgacattgccaaagatatcccaac 660
Db 601 cactgtctggggagggtccaccattgacagcagctatgacattgccaaagatatcccaac 660
Qy 661 acctgatttcattagcatcatgactacagattttctatgctgcccctggcgtgggacacag 720
Db 661 acctgatttcattagcatcatgactacagattttctatgctgcccctggcgtgggacacag 720
Qy 721 gccatcacagctccctcaggcaggtcaggagagtgcaagctcctgacagattcagcaaca 780
Db 721 gccatcacagctccctcaggcaggtcaggagagtgcaagctcctgacagattcagcaaca 780
Qy 781 ctgactatgctgtggggtacatgtttgaggtgggggctcctgcccagtaagctggtagtg 840

Db 781 ctgactatgctgtggggtacatgttgagctgggggctcctgcccagtaagctggtagtg 840
Qy 841 qcattcccacacttcgggagagagcttcaactctgcttcttctgagactggtgttcagcgc 900
Db 841 qcattcccacacttcgggagagagcttcaactctgcttcttctgagactggtgttcagcgc 900
Qy 901 caatctcaggaccgggaattccaggccggttccaaagagagcaggacccttgctact 960
Db 901 caatctcaggaccgggaattccaggccggttccaaagagagcaggacccttgctact 960
Qy 961 atgagatctgtgaacttctcccgagagccacagctccatagaacctcgccagcaggtcc 1020
Db 961 atgagatctgtgaacttctcccgagagccacagctccatagaacctcgccagcaggtcc 1020
Qy 1021 cctatgccaccaaggcgaaccagtggttaggatacagcagcaggagaaacggtcaaaaagca 1080
Db 1021 cctatgccaccaaggcgaaccagtggttaggatacagcagcaggagaaacggtcaaaaagca 1080
Qy 1081 aggtgcagctacctgaaggatagcagctggcagcgcctatggtatggccctggacctgg 1140
Db 1081 aggtgcagctacctgaaggatagcagctggcagcgcctatggtatggccctggacctgg 1140
Qy 1141 atgacttccagggtcctcttctgcgcccagatctgcgcttccctctcaacaaatgccatca 1200
Db 1141 atgacttccagggtcctcttctgcgcccagatctgcgcttccctctcaacaaatgccatca 1200
Qy 1201 aggatgcactcgtgcaacgtagcctctgttctgcacacagcagcagggggccaaagatgc 1260
Db 1201 aggatgcactcgtgcaacgtagcctctgttctgcacacagcagcagggggccaaagatgc 1260
Qy 1261 cccgttcccgtctggctggccggagcctgatacactgcctgctgagtcctcaggtgag 1320
Db 1261 cccgttcccgtctggctggccggagcctgatacactgcctgctgagtcctcaggtgag 1320
Qy 1321 cctcagctcctcctccttgggctatgcagaggttcacacacacagatttgagctcag 1380
Db 1321 cctcagctcctcctccttgggctatgcagaggttcacacacacagatttgagctcag 1380
Qy 1381 cctctgggtggcagagaggtacacacttggtagatttaattggaaatgtttacagatcccc 1440
Db 1381 cctctgggtggcagagaggtacacacttggtagatttaattggaaatgtttacagatcccc 1440
Qy 1441 aagcctggcgaaggaaatttttcaactcctgcctgcctccttctatacaggagacac 1500
Db 1441 aagcctggcgaaggaaatttttcaactcctgcctgcctccttctatacaggagacac 1500
Qy 1501 cattttggcagctctatcaccaaggagcgaacacatctctcaacacacagtgacctact 1560
Db 1501 cattttggcagctctatcaccaaggagcgaacacatctctcaacacacagtgacctact 1560
Qy 1561 aattataccccctgcaagcagcttgaaaccttcaacttaggaaacgtaattgttccct 1620
Db 1561 aattataccccctgcaagcagcttgaaaccttcaacttaggaaacgtaattgttccct 1620
Qy 1621 atctacttcccccttcttaattccacagctgctcaataaagtacaagagtttaacagtg 1680
Db 1621 atctacttcccccttcttaattccacagctgctcaataaagtacaagagtttaacagtg 1680
Qy 1681 g 1681
Db 1681 g 1681

RESULT 2
AAZ94901
ID AAZ94901 standard; cDNA; 1681 BP.
XX
AC AAZ94901;
XX
DT 01-AUG-2000 (first entry)
XX
DE Human cancer marker YKL-40 cDNA.

XX YKL-40: human; lung cancer; bronchus cancer; colorectal cancer;
KW prostate cancer; breast cancer; pancreas cancer; stomach cancer;
KW ovary cancer; bladder cancer; brain cancer; oesophagus cancer;
KW cervix cancer; melanoma; uterine endometrial cancer;
KW oral cavity cancer; pharynx cancer; liver cancer; kidney cancer;
KW biliary tract cancer; small bowel cancer; appendix cancer;
KW salivary gland cancer; thyroid gland cancer; testis cancer;
KW adrenal gland cancer; osteosarcoma; chondrosarcoma; liposarcoma;
KW malignant fibrous histiocytoma; Infection; pneumonia; meningitis;
KW arthritis; rheumatoid arthritis; osteoarthritis; fibrosis;
KW liver cirrhosis; marker; diagnosis; prognosis; ss.
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT mat_peptide 135..1223
FT /*tag= a
XX
PN WO200019206-A1.
XX
PD 06-APR-2000.
XX
XX 29-SEP-1999; 99WO-US22615.
XX
PR 01-OCT-1998; 98US-0164862.
XX
PA (RECC) UNIV CALIFORNIA.
XX
PI Price PA, Johansen JS;
XX
XX WPI: 2000-303485/26.
XX
XX
XX Novel methods for detecting cancers and evaluating the prognosis of
PT cancer using YKL-40 as a marker of cancer -
XX
XX Disclosure: Page 103-104; 111pp; English.
XX
XX The present sequence is that of the coding region of cDNA for
CC human YKL-40 mature polypeptide. YKL-40 is a 40 kDa protein
CC having Tyr, Lys and Leu as its N-terminal residues (hence, YKL-40).
CC It can be obtained from osteosarcoma cell line Mg63. YKL-40 is a
CC mammalian member of the chitinase family. It is suggested that
CC YKL-40 degrades the polysaccharide components in connective tissue
CC and/or is a lectin that binds to specific glycan structures in the
CC extracellular environment of cells. YKL-40 is useful as a marker
CC for the presence or absence of a cancer and for the prognosis of a
CC cancer. A claimed method for estimating survival length of patient
CC patients comprises obtaining a biological sample from the patient
CC and measuring the level of YKL-40, a higher level than in healthy
CC humans being indicative of reduced survival expectancy. The
CC biological sample is obtained from a cancer patient having at least
CC a preliminary diagnosis of cancer selected from lung, bronchus,
CC colorectal, prostate, breast, pancreas, stomach, ovary, urinary
CC bladder, brain, central nervous system, peripheral nervous system,
CC oesophagus, cervix, melanoma, uterine endometrial, oral cavity,
CC pharynx, liver, kidney, biliary tract, small bowel, appendix,
CC salivary gland, thyroid gland, testes, or adrenal gland cancer, or
CC osteosarcoma, chondrosarcoma, liposarcoma, or malignant fibrous
CC histiocytoma. Levels of the YKL-40 marker are elevated in
CC pathologies associated with tissue remodeling, e.g. degenerative
CC bone diseases such as rheumatoid arthritis, osteoarthritis, fibrosis,
CC cirrhosis of the liver, and cancer, especially breast, colon,
CC prostate, or lung cancer. The marker can be used to identify high
CC risk patients, and so allow selection of appropriate therapeutic
CC regimens. The methods may also be used to detect bacterial
CC infections, such as bacterial pneumonia and meningitis, as these
CC cause an elevation in YKL-40 levels, as well as diseases
CC characterized by macrophage activation, e.g. giant cell arteritis.
CC The YKL-40 marker may also be used to evaluate treatment efficacy,
CC to check for recurrence of a cancer, to monitor terminal phase
CC patients, and to check the efficacy of surgical removal of a
CC primary tumor. The methods allow estimation of the survival time

CC of patients with cancers, especially prostate, lung or colorectal
CC cancer, where the colorectal cancer is Duke's stage A, B, C, or D.
XX
SQ Sequence 1681 BP; 401 A; 503 C; 420 G; 357 T; 0 other;

Query Match 99.4%; Score 1671.4; DB 21; Length 1681;
Best Local Similarity 99.6%; Pred No. 0;
Matches 1675; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 ctaggtagctggcaccagagccgtgggcaagggaagagccacacctgctgtctg 60
Db 1 ctaggtagctggcaccagagccgtgggcaagggaagagccacacctgctgtg 60

QY 61 ctgcagccagaatgggtgtgaaggcgctctcaaacaggctttgtgctcctgctctcc 120
Db 61 ctgcagccagaatgggtgtgaaggcgctctcaaacaggctttgtgctcctgctctcc 120

QY 121 agtgcgtctctgcatacaaaactggtctgtactacaccagctggtccccagtcgggaag 180
Db 121 agtgcgtctctgcatacaaaactggtctgtactacaccagctggtccccagtcgggaag 180

QY 181 gcgatgggagctgcttcccagatgccccttgacccgtctctctgtgtaccacatcatctaca 240
Db 181 gcgatgggagctgcttcccagatgccccttgacccgtctctctgtgtaccacatcatctaca 240

QY 241 gctttgccaatataagcaacgatacacatgcagacacctgggagtggaatgatgtgacgctct 300
Db 241 gctttgccaatataagcaacgatacacatgcagacacctgggagtggaatgatgtgacgctct 300

QY 301 acggcatgctcaacacacitcaacaacacgaaccccaacccctgaagactctcttctgtctcg 360
Db 301 acggcatgctcaacacacitcaacaacacgaaccccaacccctgaagactctcttctgtctcg 360

QY 361 gaggatggaactttgggtctcaagattttccaaagatagcctcccaacacccagatgcgc 420
Db 361 gaggatggaactttgggtctcaagattttccaaagatagcctcccaacacccagatgcgc 420

QY 421 gaactttcatcaagtacgtaccgcctattttctgcgaccctatggcttttgatggcgctgacc 480
Db 421 gaactttcatcaagtacgtaccgcctattttctgcgaccctatggcttttgatggcgctgacc 480

QY 481 ttgctctgctctaccctggacggagagacaaacacacattttaccaccttaatacaaggaaa 540
Db 481 ttgctctgctctaccctggacggagagacaaacacacattttaccaccttaatacaaggaaa 540

QY 541 tgaaggccgaatttataaagggaagccagccagccagggaagagcagctctctcagcgag 600
Db 541 tgaaggccgaatttataaagggaagccagccagccagggaagagcagctctctcagcgag 600

QY 601 cactgtctcgggggaagggtccaccattgacagcagctatgacattgccaagatattcccaac 660
Db 601 cactgtctcgggggaagggtccaccattgacagcagctatgacattgccaagatattcccaac 660

QY 661 acctgattcttaqcatcatgacacacacattttcatggcctgctggcgctgggaccacag 720
Db 661 acctgattcttaqcatcatgacacacacattttcatggcctgctggcgctgggaccacag 720

QY 721 gccatcacagtcctccctcagcgaggtcagagagatgcaagctctgcagagattcagcaaca 780
Db 721 gccatcacagtcctccctcagcgaggtcagagagatgcaagctctgcagagattcagcaaca 780

QY 781 ctgactatgctgtggggtacatgtttgagctgggggtcctctgcagtaagctggtgatgg 840
Db 781 ctgactatgctgtggggtacatgtttgagctgggggtcctctgcagtaagctggtgatgg 840

QY 841 gcatacccaccttcgsggagagagcttcactctgctctctcttgagactggtgttccagcgc 900
Db 841 gcatacccaccttcgsggagagagcttcactctgctctctcttgagactggtgttccagcgc 900

QY 901 caatctcaggaccgcgaattccagccggttcaccaggagcgaggaccccttgcctact 960
Db 901 caatctcaggaccgcggaattccagccggttcaccaggagcgaggaccccttgcctact 960


```

Db 661 cccctgtccgaggtcaggaggtgcaagtccctgcacagattcagcaaacact----- 711
QY 792 gtgggtatcatgttgaggtctggggctcctgccagtaagctggtgatggcctccacc 851
Db 712 ----- 711
QY 852 ttcgggaggtctcactctggtcttcttgagactggtgttccacgccaatctcagga 911
Db 712 ----- 711
QY 912 ccgggaattccagccggttccaccaaggaggcaggacccttgctactatagatctgt 971
Db 712 -----atctgt 717
QY 972 gaattctccgagagaccacagtcacatagaaccctcgccagcaggtccctatgccacc 1031
Db 718 gaattctccgagagaccacagtcacatagaaccctcgccagcaggtccctatgccacc 777
QY 1032 aaggccaaccagtggttaggatacagcagcaggaagcgtcaaaagcgaaggtgcagtac 1091
Db 778 aaggccaaccagtggttaggatacagcagcaggaagcgtcaaaagcgaaggtgcagtac 837
QY 1092 ctgaagataggcagctggcagcgccatggtatggccctggacctggatgacttccag 1151
Db 838 ctgaagataggcagctggcagcgccatggtatggccctggacctggatgacttccag 897
QY 1152 ggtcctctcgccagcagctggtcctcctcctcctcctcctcctcctcctcctcctc 1211
Db 898 ggtcctctcgccagcagctggtcctcctcctcctcctcctcctcctcctcctcctc 957
QY 1212 gctgcaacg 1220
Db 958 gctgcaacg 966

RESULT 5
AAC00233
ID AAC00233 standard; cDNA; 503 BP.
AC AAC00233;
XX
DT 06-OCT-2000 (first entry)
DE Human secreted protein 5' EST, SEQ ID NO: 231.
KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KW gene therapy; chromosome mapping; ss.
XX Homo sapiens.
XX EP1033401-A2.
XX 06-SEP-2000.
XX 21-FEB-2000; 2000EP-0200610.
XX 26-FEB-1999; 99US-0124287.
XX (GEST ) GENSET.
XX Dumas Milne Edwards J, Duclert A, Giordano J;
XX WPI; 2000-500381/45.
XX P-PSDB; AAG00227.
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 1; SEQ ID 231; 71pp + CD-ROM; English.
XX PS
XX CC The present sequence is one of a large number of 5' ESTs derived from
XX mRNAs encoding secreted proteins. An ORF has been identified within the

```

```

CC sequence. The 5' ESTs were prepared from total human RNAs or polyA+ RNAs
CC derived from 30 different tissues. EST sequences usually correspond
CC mainly to the 3' untranslated region (UTR) of the mRNA because they are
CC often obtained from oligo-dT primed cDNA libraries. Such ESTs are not
CC well suited for isolating cDNA sequences derived from the 5' ends of
CC mRNAs and even in those cases where longer cDNA sequences have been
CC obtained, the full 5' UTR is rarely included. 5' ESTs are derived from
CC mRNAs with intact 5' ends and can therefore be used to obtain full length
CC cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic,
CC gene therapy and chromosome mapping procedures. They are used to obtain
CC upstream regulatory sequences and to design expression and secretion
CC vectors.
XX
SQ Sequence 503 BP; 121 A; 140 C; 135 G; 107 T; 0 other;

Query Match 26.5%; Score 446.2; DB 21; Length 503;
Best Local Similarity 99.3%; Pred. No. 4.4e-117;
Matches 448; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 ctaggtagctggcaccagagccgtggcagggaagggaagggccacacccctgcctgtctg 60
Db 53 ctaggtagctggcaccagagccgtggcagggaagggaagggccacacccctgcctgtctg 112
QY 61 ctgagccagaatgggtgtgaaggcgtctcaaacagagcttggctcctagtgctactcc 120
Db 113 ctgagccagaatgggtgtgaaggcgtctcaaacagagcttggctcctagtgctactcc 172
QY 121 agtgcgtctctgcatacaaaactggtctgtactacacacagctgtccacagtcgggaag 180
Db 173 agtgcgtctctgcatacaaaactggtctgtactacacacagctgtccacagtcgggaag 232
QY 181 gcatgggagctgcttccagatgcccctgacgcttccctgacgcttccctgtaccacatcata 240
Db 233 gcatgggagctgcttccagatgcccctgacgcttccctgacgcttccctgtaccacatcata 292
QY 241 gcttgcacaataaagcaacagatcacatgacacacccctggagtggaatgatgacgctct 300
Db 293 gcttgcacaataaagcaacagatcacatgacacacccctggagtggaatgatgacgctct 352
QY 301 acgcatgctcaacacactcaacacacacgacaccccaacctgaagactctctgtctgtcg 360
Db 353 acgcatgctcaacacactcaacacacacgacaccccaacctgaagactctctgtctgtcg 412
QY 361 gaggatggaacttgggtctcaagatttccaaagatagcctccacacacacagatgagcc 420
Db 413 gaggatggaacttgggtctcaagatttccaaagatagcctccacacacacagatgagcc 472
QY 421 ggactttcatcaagtacgtaccgcccatttct 451
Db 473 ggactttcatcaagtacgtaccgcccatttct 503

RESULT 6
AAT50834
ID AAT50834 standard; cDNA; 1713 BP.
XX
AC AAT50834;
XX
DT 24-MAR-1997 (first entry)
DE Human chitinase cDNA clone chi.39.
XX
KW Chitinase; chitotriosidase; chitin; infectious disease;
KW gene therapy; vaccine; lysosomal lipidosis; Gaucher's disease;
KW leishmaniasis; sarcoidosis; X-linked adrenoleukodystrophy;
KW multiple sclerosis; drug delivery; cosmetics; food; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 13..1176
XX /*tag= a
XX FT

```



```
FT sig_peptide /product= "Human chitinase from clone pMO-218"  
FT 2..64 /*tag= b  
FT mat_peptide 65..1399  
FT /*tag= c  
FT /product= "Human mature chitinase from clone pMO-218"  
FT W0200123430-A2.  
PN  
XX  
XX  
PD 05-APR-2001.  
XX  
XX 28-SEP-2000; 2000MO-US26960.  
PF  
XX 30-SEP-1999; 99US-0409918.  
PR  
XX (ICOS-) ICOS CORP.  
PA  
XX  
PI Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;  
XX  
XX WPI; 2001-266141/27.  
DR P-PSDB; AAE00432.  
DR  
XX Novel chitinase immunoglobulin fusion product, useful for treating  
PT fungal infections and reducing the amount of a non-chitinase antifungal  
PT agent needed for the treatment.  
PT  
XX Claim 2; Page 29-31; 39pp; English.  
XX  
CC The present invention relates to a chitinase immunoglobulin (Ig) fusion  
CC product, comprising a human chitinase fused to at least a portion of an  
CC immunoglobulin chain. The fusion product is useful for treating fungal  
CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,  
CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,  
CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.  
CC The fusion protein is useful for reducing the amount of non-chitinase  
CC antifungal agent needed to exert an antifungal activity. The fusion  
CC protein is also useful for preparing a medicament for the prophylactic  
CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin  
CC fusion product has unexpectedly improved serum half-life and formulation  
CC properties. The present sequence is human chitinase cDNA from clone  
CC pMO-218. Chitinase enzyme degrades chitin which is a homopolymer of  
CC beta-(1,4)-linked N-acetylglucosamine residues.  
XX  
SQ Sequence 1636 BP; 361 A; 491 C; 440 G; 344 T; 0 other;  
  
Query Match 23.6%; Score 397.2; DB 22; Length 1636;  
Best Local Similarity 62.9%; Pred. NO. 6.9e-103;  
Matches 691; Conservative 0; Mismatches 383; Indels 24; Gaps 4;  
  
QY 94 caggcttggctcctgggtgctgctcagtgctgctgcatacaaacctggtgctgact 153  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
24 caggcttgcagtcctgctgctgctgctcctgctgctgctgctgctgctgctgctgct 83  
  
QY 154 acaccagtgctcctcagtcaggagagagagagagagagagagagagagagagagag 213  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
84 tcacacactgggcccagtcacagacagagagagagagagagagagagagagagagagag 143  
  
QY 214 gcttcctgtgtaccacacacacacacacacacacacacacacacacacacacacacac 273  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
144 ccagccttgcacccacacacacacacacacacacacacacacacacacacacacacac 203  
  
QY 274 cctggagtggaatgatgagcctcctcagcagcagcagcagcagcagcagcagcagcag 333  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
204 ccactgagtggaatgagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 263  
  
QY 334 ccaacctgaagactctctgctgctcggagagagagagagagagagagagagagagagag 393  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
264 ccaagctgaagacccctgttagccatcgaggagagagagagagagagagagagagagag 323  
  
QY 394 agatagcctcacaacacacacagagtcgcggagccttcacagtcagtcagtcagtcagtc 453  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
324 atatgtagccacgcccacaacacagtcagacacacacacacacacacacacacacacac 383
```

```
QY 454 gcaaccatggttgatggcgtagccttgacctgctcctggtgcttaccctgacgag----- 505  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
384 gcaatacagcttgcgcttgcgcttgccttgccttgccttgccttgccttgccttgccttgc 443  
  
QY 506 -----agacaacacacacattttaccacacacacacacacacacacacacacacacac 558  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
444 ctgctgtagacaaagagcgcttcacacacacacacacacacacacacacacacacacacac 503  
  
QY 559 aggaagccacag-----ccagggaaaaagcagcctcctgctcagcgagcagcagcagcagcag 615  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
504 aggaagccacagcctcagcaggaagcagccttctcttctgagtgagcagcagcagcagcagc 563  
  
QY 616 aggtcacattgacacagcagctatgacatgccaagatatcccaacacacacacacacacacac 675  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
564 agacctatgtggtatgtagacgaggtggacaaaatcgccacagacacacacacacacacac 623  
  
QY 676 gcatcatgacctacgatttttcgctgctgctgctgctgctgctgctgctgctgctgctgct 735  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
624 acctatgacctacgacttccatgctcttctggaaggtcacggaacacacacacacacacac 683  
  
QY 736 tcaggcaggtcagcagcaggtgcaagtcctgacagattcagcaacacacacacacacacacacac 795  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
684 tctacaagagggcagaagagagtggtgcagcagccagcctcaacgtggatgctgctgtgc 743  
  
QY 796 ggtacatgttgagcctggggcctcctgccaagtaagcgtggtgagtgagcagcagcagcagc 855  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
744 aacagtggctgcaagaggggacccctgccaagcagctgacctggcctgcttaccctacac 803  
  
QY 856 ggaggagcttcaactcctg-----cttcttctgagactggtgttccagcgccacacacacacac 912  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
804 gacgctccttcacacgctcctcctcctcctcctcctcctcctcctcctcctcctcctcctc 863  
  
QY 913 cgggaattccagcggccttccacaaagagcagcagcagcagcagcagcagcagcagcagcagc 972  
DB ctggcactccagcggccttccacaaagagcagcagcagcagcagcagcagcagcagcagcagc 921  
  
QY 973 actctcctcgcgagccacagctccatagacacctcgccagcagcagcagcagcagcagcagc 1032  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
922 -ctcctggagggggggccacaaacagagaatccagcagcagcagcagcagcagcagcagcagc 980  
  
QY 1033 agggcaaccagctgggtagatagcagcagcagcagcagcagcagcagcagcagcagcagcag 1092  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
981 gggacaaccagctgggtgggcttggatgagtgagagagcttcaaaacacacacacacacac 1040  
  
QY 1093 tgaagatagcagcagctggcagcagcagcagcagcagcagcagcagcagcagcagcagcag 1152  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
1041 tgaagcagaagggagcaggcgggcagcagcagcagcagcagcagcagcagcagcagcagcag 1100  
  
QY 1153 gctccttctgcggccagg 1170  
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
1101 gcttctcctgcaaccagg 1118  
  
RESULT 11  
AAT89181  
ID AAT89181 standard; cDNA; 1768 BP.  
XX  
AC AAT89181;  
DT XX  
DT 27-APR-1998 (first entry)  
XX  
DE Human chitotriosidase variant cDNA.  
XX  
XX Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;  
KW rheumatoid arthritis; atherosclerosis; human; ss.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 124..1524  
FT CDS /*tag= a  
FT  
XX
```


FT /note= "from clone MO-13B"
 PN W09747752-A1.
 XX 18-DEC-1997.
 PD 16-JUN-1997; 97WO-US10460.
 XX 14-JUN-1996; 96US-0663618.
 XX (ICOS-) ICOS CORP.
 PA Gray PW;
 PI WPI; 1998-052316/05.
 DR P-PSDB; AAW40260.
 XX Nucleic acids encoding human chitinase - useful as antifungal
 PT agents, especially in combination with other antifungals
 XX Claim 9; Page 42-44; 63pp; English.
 CC This sequence encodes a novel human chitinase isolated from clone MO-13B.
 CC Chitinases are useful for treating or preventing fungal infection and
 CC as immunogens for generating antibodies which are used to purify, detect
 CC and quantify chitinases, e.g. for diagnosis of Gaucher's disease. The
 CC nucleic acid sequence of the chitinase is also useful as a probe to
 CC identify and isolate genomic DNA encoding chitinases or similar proteins,
 CC or cells expressing them or to generate transgenic ('knockout') rodents.
 CC It can also be used in hybridisation assays and to detect genetic
 CC alterations in the chitinase gene related to disease. Agents that inhibit
 CC this protein may be useful in treatment of Gaucher's disease and
 CC rheumatoid arthritis, where overexpression of the protein can damage
 CC the extracellular matrix. Chitinase also improves the activity of other
 CC antifungal agents and may allow a reduction in the dose of such agents,
 CC and thus of their side effects.
 XX Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;
 SQ

Query Match 23.6%; Score 396; DB 19; Length 1656;
 Best Local Similarity 62.3%; Pred. No. 1.5e-102;
 Matches 700; Conservative 0; Mismatches 400; Indels 24; Gaps 4;

QY 68 cagaatgggtgaagcgtctcaaacagcgtttgtggctctgtgctgctcagtgctg 127
 DB 23 catcatgggtgcgtgtggcctggcaggtttcatggctcgtgctgctgctcagtgctg 82

QY 128 ctctgcatacaactggtctgtactacacacagctggtccagtagcgggaagcgatgg 187
 DB 83 ctctgtgcaaaactggtctgtacttccaaactggcccaactgagcagagagggaggc 142

QY 188 gagctgtctccagatgccccttgacgcttctctgtatccacatcatctacagctttgc 247
 DB 143 tgccttctgcccgaagacttggaaccagccttggtgcacccctcatctacgcttcgc 202

QY 248 caataaagaacagatcacatgacacacctggagtggaatgatgtgacgtctacggcat 307
 DB 203 tggcatgaccacacacagctgagcaccactgagtggatgacgagactctctaccagga 262

QY 308 gctcaacacactcaacaacagaccccaacctgagagactctctgtctgctgagagatg 367
 DB 263 gtcaatggcctgaagaagatgaatcccaagctgagagacctgttagcctatcgaggctg 322

QY 368 gaacttgggtctcaagaattttccaagatagcctcccaacacacacagtagtcgcggacttt 427
 DB 323 gaatttcagactcagaagttcacagatatgtgccaagggcccaacacgcgtcagacctt 382

QY 428 catcaagtcagtaaccgccattttctgcgaccccatgctttgatggcgtgacacttgcctg 487
 DB 383 tgtcaactggcctcaggtttctgcgcaaatacagctttgacggccttgaccttgactg 442

QY 488 gctctaccctggagc-----gagagacaaacacacattttaccaccctaatt 532

DB 443 ggagtaccacaggaagccaggggagccctgcgtgagacaagagcgtcttcacaacctggt 502
 QY 533 caaggaaatgaaggccgaattttataaagaagcccaag---ccagggaataaagcagctct 589
 DB 503 acaggacttggcgaatgcctctccagcagggaagcccaagcctcagggaaggaacgccttct 562
 QY 590 gctcagcgcagcactgtctgcggggaaggtcacccattgacagcagctatgacattgccaa 649
 DB 563 tctgagtgcagcgggttccagctggcgagacctatgtggatgctgatacagagtgacaa 622
 QY 650 gatatcccaacctggatttcatatgacatcatgacactacgattttcatggcgctcgcg 709
 DB 623 aatgcgccagaacctggtttgttcaaccttatggctacgacttccatggctcttggga 682
 QY 710 tgggaccacagcccatcacagctccctcaggcgaggtcaggagatgcaagtcctctacag 769
 DB 683 gaaggtcacgggacataacagccccctctacaagaggaagagagagtggtgcagcagc 742
 QY 770 attcagcaacactgactatgtgtgggtacatgtgaggtcgtggggctcctgccagtaa 829
 DB 743 cagctcaacgtggatgctgtgtgcaacagtggtgcagaaggggagccctgccagcaa 802
 QY 830 gctggtgatgggtacccacacttctcgaggagagcttcaactctgg---cttcttctgagac 886
 DB 803 gctgactcttggcatgcttacctacgcagcgtctctcactggtcctctcatcagacac 862
 QY 887 tgggttccagcgcacaaatctcagggaccgggaattccagcccggttccacaagagcagc 946
 DB 863 cagagtggggggcccccagcacaggtgtcgtcactcagggcccttccaccaagaaagg 922
 QY 947 gaccttgcctactaagagatctgtgacttctctccgcgagccacagtcctatagaacct 1006
 DB 923 gatgctggcctactatgaagtctg---ctcctggaagggggccaccaacacagagaatcca 979
 QY 1007 cgccagcagctcccttatgtccaccaaggggcaaccagtggtgtaggatacagcagcagga 1066
 DB 980 ggatcagaaggtgccctacatcttccgggacaaccagtggtggcttgatgtgga 1039
 QY 1067 aagcgtcaaaagcaagtgtagctacctgacagtagagtagagcgtggcgagcagcagtgatg 1126
 DB 1040 gagcttcaaaacaaagtcagctatctgaagcagaagggactggcggggccaatggtctg 1099

QY 1127 ggccttggacctgagacttccagggctcctcttctcgccagg 1170
 DB 1100 ggcactggacttagacttggcgggtctctctcgaaccagg 1143

RESULT 13
 AA221848
 ID AA221848 standard; DNA; 1656 BP.
 XX
 AC AA221848;
 XX
 DT 10-DEC-1999 (first entry)
 XX
 DE MO-13B clone of human Chitinase, with noncoding 5'/3' regions.
 XX
 KW chitin; fungal infection; immunocompromised; AIDS; chemotherapy;
 KW organ transplant; parasite; chitin-binding; allele; vector;
 KW truncated protein; ds.
 XX
 OS Homo sapiens.
 XX
 Key Location/Qualifiers
 FT 27..1427
 CDS /tag= a
 FT /product= Human_Chitinase
 FT 27..89
 sig_peptide /tag= b
 FT /note= "Signal peptide"
 FT 90..1427
 mat_peptide /tag= c

/note= "Mature peptide"

W09946390-A1.

16-SEP-1999.

12-MAR-1999; 99WO-US05343.

12-MAR-1998; 98US-0039198.

(ICOS-) ICOS CORP.

Gray PW, Tjoelker LW;

WPI; 1999-551417/46.

P-PSDB; AAY42426.

Novel chitin-binding fragments of human chitinase used to treat fungal infections in animals

Example 1; Page 59-62; 83pp; English.

This is the nucleotide sequence of an allelic form of the human chitinase enzyme, which is capable of degrading chitin (a linear homo polymer of beta-1,4-linked N-acetylglucosamine residues). Chitinase fragments can be used to screen for proteins or other molecules that specifically bind to the chitin-binding domain of human chitinase or that modulate its activity. These compounds are useful for immunization, as well as for purifying chitinase, as well as for detection and quantification of chitinase. Polynucleotide fragments of the invention are useful as a source of probes and primers, and to express the proteins recombinantly. The chitinase fragments, when conjugated to antifungal compounds, are used to treat animals, especially humans, infected with chitin-containing parasites such as fungi. Fungal infection treated include candidiasis, aspergillosis, mucormycosis, histoplasmosis, paracoccidioidomycosis, sporotrichosis, and dermatophytoses. Use of whole chitinase protein for treating infections, especially fungal infections, is problematic. In view of the increasing incidents of life-threatening fungal infection in e.g. immunocompromised individuals, there exists a need for identifying new compounds for treating fungal infection. The chitin-binding fragments of the present invention provide this need.

Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Query Match 23.6%; Score 396; DB 20; Length 1656;

Best Local Similarity 62.3%; Pred. NO. 1.5e-102;

Matches 700; Conservative 0; Mismatches 400; Indels 24; Gaps 4;

QY 68 cagaatgggtggaaggcgtctcaacagcgtttgtggtccctgggtgctgctcagtgctg 127
 DB 23 catcatgggtgctgtgctggcgtggcaggtttcattgttctcgtgctgagatccatggg 82
 QY 128 cctgcatacaactggtctgctactacacagcgtggtccagtagtaccgggaaggcgtg 187
 DB 83 cctcgtgcaaaactggtctgctacttcccaactggccagtagacagacaggggagc 142
 QY 188 gagctgttcccaagtcgccctgagcgttctctgtgtaccacatcatctacagcttgc 247
 DB 143 tcgcttctgccaaggacttggaacccagcgtttgacccaccctcatctacgcttcgc 202
 QY 248 caataaagcaacgatacacatgcacacctgggagtggaatgatgtgacgtctctacggcat 307
 DB 203 tggcatgaccaacacacagctgagccactgagtggaatgacgagactctaccagga 262
 QY 308 gctcaacacactcaacacacgacacccacacctgaagactctcttctgtgtggaggtg 367
 DB 263 gttcaatggcctgaagaagatgaatcccaagctgaagaccctgttagccatcgaggctg 322
 QY 368 gaacttgggtctcaagaattttccaagatagcgtcccaacacccagagtcgcggacttt 427

RESULT 14

AAD03760

ID AAD03760 standard; cDNA; 1656 BP.

XX AAD03760;

XX 19-JUN-2001 (first entry)

XX Human chitinase cDNA from clone pMO-13B.

XX Human; antifungal; chitinase; immunoglobulin; Ig; therapy;

KW fungal infection; candidiasis; aspergillosis; coccidioidomycosis;

KW blastomycosis; paracoccidioidomycosis; histoplasmosis; cryptococcosis;

KW chromoblastomycosis; sporotrichosis; mucormycosis; dermatophytoses;

XX clone pMO-13B; ss.

XX Homo sapiens.

OS

Db 323 gaatttcagcactcagaagttccagatatggttagccacgcccacacacgctcagacctt 382
 QY 428 catcaagtacgtaccgcacatttctgcgcacccattggttgatggcggtgaccttgcgtg 487
 Db 383 tgtcaactcgccatcaggtttctgcgcaaatatacagctttgacggccttgaccttgactg 442
 QY 488 gctctaccctggacg-----gagagacaaacacacattttaccacccctaat 532
 Db 443 ggagtaccacaggaagccaggggagccctgcggtagacaagagagcgcttcacaaacctggt 502
 QY 533 caaggaaatgaaggccgaattataaaggaaagccag---ccagggaagaaagcgactcct 589
 Db 503 acaggacttgccaatgctctccagcaggaagccacacctcagggaaggaagcgccttct 562
 QY 590 gctcagcgcacactgtgtggtggggaaggtcacatttgacagcagcagctatgacattgcca 649
 Db 563 tctgagtcagcgggttcagctgggcagacctatgtggtgctggtgacacaggggggacaa 622
 QY 650 gatatacccaaacactggatttcatttagcatcatcatttttcatggtgcgcctggcg 709
 Db 623 aatcgcccaagaacctggattttgtcaaccttatgctcagcacttccatggtcttggga 682
 QY 710 tgggacccacagccatcacagtcctcctcagcgaggtgcaggaggtgcaagtcctcgacac 769
 Db 683 gaaggtcacgggacataacacagccctctacaaagagcgaagagagtggtgcagcagc 742
 QY 770 attcagcaacactgactatgctgtgggtggtacattgttgagctgggggctcctgcagtaa 829
 Db 743 cagcctcaacgtggatgctgtgcaacagtggtgcagaggggggacccctgcagcaa 802
 QY 830 gctggtgagggcattcccccactcctgggaggttcactctggt---cttctcttgagac 886
 Db 803 gctgactcttggtatgctcctacacagcgtccttcacactggcctcctcatcagacac 862
 QY 887 tgggtttccagcggcacaatctcagacccgggaattccagccggttcccaagagcagc 946
 Db 863 cagagtgggggcccagccacaggtgtgctgactcagccctccacaaaggaaggagg 922
 QY 947 gaccttgctactatgagatctgacttcctccgagcagccacagtcacatagaacctt 1006
 Db 923 gatgtggtcctactatgaagtctg---ctcctggaagggggccaccaaagagaatcca 979
 QY 1007 cggccagcaggtccctctatgccaccaggggcaaccagtggtggtgagatacagacacgga 1066
 Db 980 ggatcagaaggtgctcactatctccgggacacacagtggtggtggttggatgatgga 1039
 QY 1067 aagcgtcaaaagcaaggtgcagtcactgaaggtagcagcagtcgagcggccatggtatg 1126
 Db 1040 gagttcaaaacacaggtcagtcactatctgaagcagaagggagctggggggccatggtctg 1099
 QY 1127 ggcctgagcctggatgacttccagggctcctctctcgcgcagcagg 1170
 Db 1100 ggcactggacttagactttgctggcgttctcctcgcacacagg 1143

```

XX FH Location/Qualifiers
FT CDS 27..1427
FT /tag= a
FT /product= "Human chitinase from clone pMO-13B"
FT /tag= b
FT /tag= c
FT /product= "Human mature chitinase from clone pMO-13B"
XX PN WO200123430-A2.
XX PD 05-APR-2001.
XX XX
XX PF 28-SEP-2000; 2000WO-US26960.
XX PR 30-SEP-1999; 99US-0409918.
XX PA (ICOS-) ICOS CORP.
XX PI Allison DS, Dietsch GN, Gray PW, Shaw KD, Steiner BH;
XX WPI; 2001-266141/27.
XX DR P-PSDB; AAE00433.
XX PT Novel chitinase immunoglobulin fusion product, useful for treating
XX PT fungal infections and reducing the amount of a non-chitinase antifungal
XX PT agent needed for the treatment -
XX PS Claim 2; Page 34-36; 39pp; English.
XX CC The present invention relates to a chitinase immunoglobulin (Ig) fusion
XX CC product, comprising a human chitinase fused to at least a portion of an
XX CC immunoglobulin chain. The fusion product is useful for treating fungal
XX CC infections (mycoses) such as candidiasis, aspergillosis, blastomycosis,
XX CC coccidioidomycosis, paracoccidioidomycosis, histoplasmosis, mucormycosis,
XX CC cryptococcosis, chromoblastomycosis, sporotrichosis and dermatophytoses.
XX CC The fusion protein is useful for reducing the amount of non-chitinase
XX CC antifungal agent needed to exert an antifungal activity. The fusion
XX CC protein is also useful for preparing a medicament for the prophylactic
XX CC or therapeutic treatment of fungal infections. Chitinase immunoglobulin
XX CC fusion product has unexpectedly improved serum half-life and formulation
XX CC properties. The present sequence is human chitinase cDNA from clone
XX CC pMO-13B. Chitinase enzyme degrades chitin which is a homopolymer of
XX CC beta-(1,4)-linked N-acetylglucosamine residues.
XX SQ Sequence 1656 BP; 365 A; 497 C; 447 G; 347 T; 0 other;

Query Match 23.6%; Score 396; DB 22; Length 1656;
Best Local Similarity 62.3%; Pred. NO. 1.5e-102;
Matches 700; Conservative 0; Mismatches 400; Indels 24; Gaps 4;

QY 68 cagaatgggtgtaaggctctcaaacagagctttgtgctcgtggtgctcctcagtgctg 127
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 23 catcaatgggtgctgtggtgctggtgaggtttcatggtctgctgctgatgccatggg 82
QY 128 ctctgatacaactggctgtactacacacagctggtgccagctaccgggaagcgatgg 187
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 83 ctctgtgcgaactggctgtacttaccacactggccacagctggtccagatgagcaggc 142
QY 188 gactgtctccagatgcccctgacgcttccgtgtatccacatcatcatcagctttgc 247
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 143 tcgtctctgccaaaggacttgacccagccttgcacccacccctcatcacccttcgc 202
QY 248 caataagaacacgatacatgacacactggagtggaatgatgtgacgtctcagggcat 307
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 203 tggcatgacacacacacagctgagcaccactgagtgaatgacgagactctcaccagga 262
QY 308 gctcaacacactcaacaacacgaccccaacctgaagactctcttctgtcggaggtg 367
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 263 gtccaatggcctgaagaagatgaatcccaagctgaagaccctgttagccatcgaggctg 322

```

```

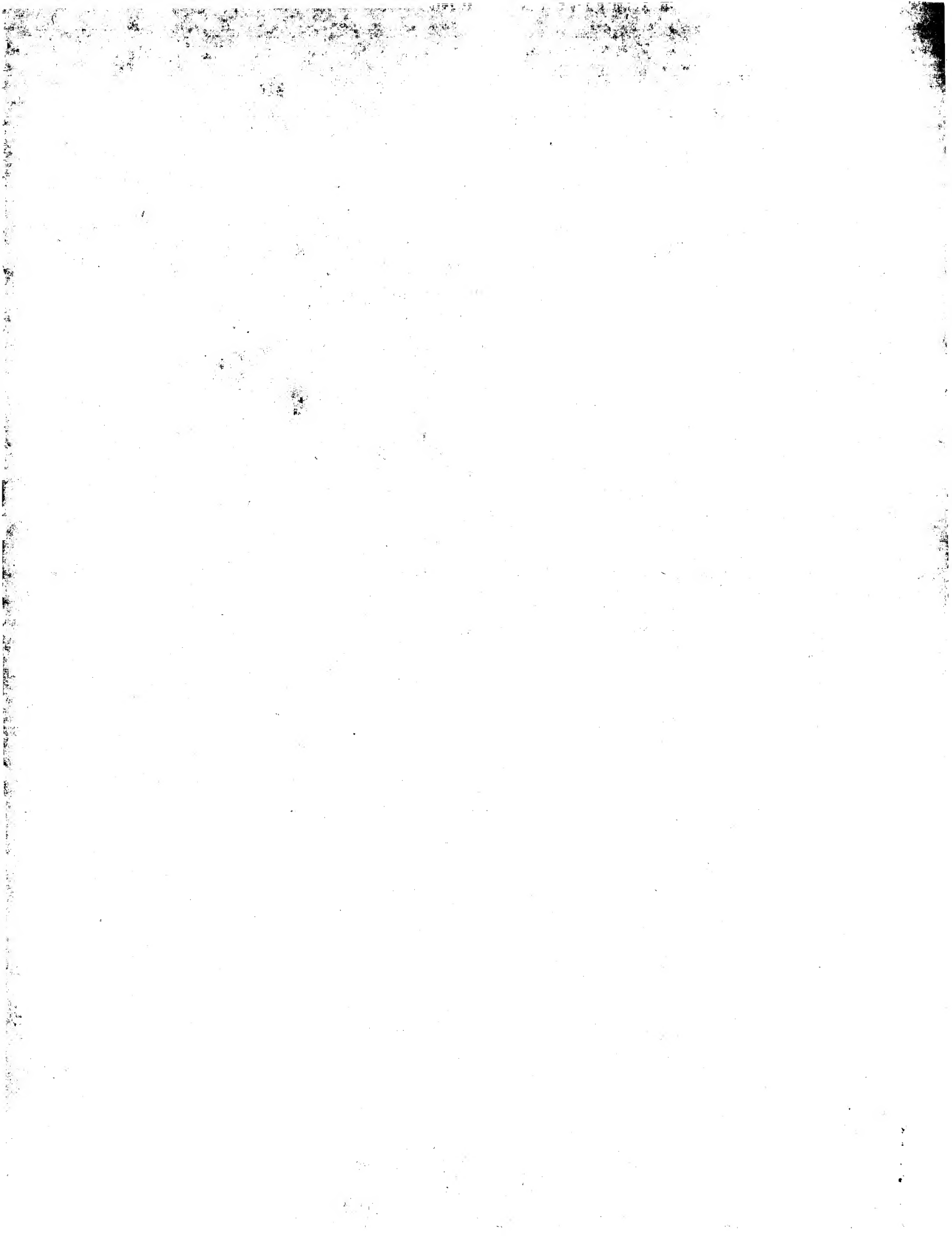
QY 368 gaactttgggtctcaagattttccaagatagcctccaacacccagagtcgcggacttt 427
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 323 gaatttcagcactcagaagttcacagatatggtagcagcgccaacaacgcgtcagacctt 382
QY 428 catcaagtcagtcaccgccatttctgcgaccccatggtcttgatgagcgctgacacctg 487
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 383 tgtcaactcggccatcaggtttctgcgcaatacagctttgacggccttgacacctgactg 442
QY 488 gctctaccctggagc-----gagagacaaacacacattttacacccctaat 532
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 443 ggagtaccacagggaagccagggggagccctgcgtagacaaggagcgcttccacacccctg 502
QY 533 caaggaaatgaagccggaatttataaagggaagccacg---ccagggaacacagcagctct 589
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 503 acaggacttgccaatgctctccagcaggaagccacagacctcaggaaggaagcagctctt 562
QY 590 gctcagcgcagcactgtctgcgggggaaggtcaaccttgacacagcagctatgacattgccaa 649
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 563 tctgagtcagcgggttccagctggcagacctatgtggtgctggatagcaggtggacaa 622
QY 650 gatattcccaacactggatttcaattagcatcatgacctaagcagattttcaatggcctggcg 709
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 623 aatcgcacgaacctgattttgtcaaccttatggcctacgacttccatggctcttgga 682
QY 710 tgggacacagccatcacacagctccctcaggcaggttcaggagagatgcaagtctctgacag 769
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 683 gaaggttcacggggcatacaacagccccctctacaagagggcaagagagagtggtgcagcagc 742
QY 770 attcagcaacactgactatgctgtgggtgtacatgtgtgaggtcgggggctcctgccagtaa 829
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 743 cagcctcaacgtggatgctgtgtgcaacagtggtgcaagaaggggacccctgccagcaa 802
QY 830 gctgggtgaggggcatccccacacttcggggagagagcttcaactctgg---cttctctgagac 886
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 803 gctgactccttggtgctctacactcagcagcgtctctcactcgtcctcctcctcactcagacac 862
QY 887 tgggtgtccagcgcacatctcaggaccgggaattccagcgcggttcaccaagagagcag 946
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 863 cagagtgggggccccagccacaggggtcggcactccagcgccttccaccaaggaagaggg 922
QY 947 gaccttgccctactatgagatctgtgacttctcctcgggagccacagtcctatagaacct 1006
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 923 gatgtggcctactatgaagctctg---ctcctggaagggggccaccaaacacagagaatcca 979
QY 1007 cgccagcaggtccctctatgcaccaagggcaacccagtggtggtggtggtggtggtggtggtg 1066
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 980 ggatcagaaggtgcttccatcttccggggaacacacagtggtggtggtggtggtggtggtg 1039
QY 1067 aagcgtcaaaagcaaggtgcagtcacctgaagagatagcagctggcagcgccatggtatg 1126
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 1040 gacttcaaaaccaaagtcagctatctgaagcagaagaggaactggggtggggtggtggtg 1099
QY 1127 ggccttggaactggatgacttccagggtcctctcttcgcggccagg 1170
DQ || || || || || || || || || || || || || || || || || || || || ||
DQ 1100 ggcactggacttagagcacttgcgggtctctctctctctgcaaccagg 1143

RESULT 15
AAT89180
ID AAT89180 standard; cDNA; 1637 BP.
XX
XX AAT89180;
XX
XX 27-APR-1998 (first entry)
XX
XX Human chitotriosidase cDNA.
DE
DE Chitotriosidase; tissue remodelling disorder; diagnosis; therapy;
KW rheumatoid arthritis; atherosclerosis; human; ss.
XX
XX Homo sapiens.
XX
XX

```


QY 1073 caaagcaaggtgcagtaacctgaaggataggcagctggcgaggcgcccatggtatgggccct 1132
||||| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
Db 1040 caaaaccaaggtcagctatctgaagcagaaggactggcgggggcccatggtctgggcact 1099
QY 1133 ggacctggatgacttccagggtcccttcttcgcggccagg 1170
||||| ||||| || ||||| || ||||| || ||||| || ||||| || ||||| ||
Db 1100 ggacttagatgactttgcgggttcttctcctgcaaccagg 1137

Search completed: December 6, 2001, 23:36:45
Job time: 7750 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2001, 22:24:55 ; Search time 93.86 Seconds

(without alignments)

4056.142 Million cell updates/sec

Title: US-09-164-862b-4

Perfect score: 1681

Sequence: 1 ctaggtagctggcaccagga.....tacaagagtttaacagtgtg 1681

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 351203 seqs, 113238999 residues

Total number of hits satisfying chosen parameters: 702406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued_Patents_NA.*

1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*

6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|--------|-------|--------|-------|-------------------|
| 1 | 1681 | 100.0 | 1681 | 2 | US-08-581-527-4 |
| 2 | 1674.6 | 99.6 | 1681 | 5 | PCT-US94-07754-4 |
| 3 | 397.6 | 23.7 | 1643 | 2 | US-08-486-839-3 |
| 4 | 397.6 | 23.7 | 1643 | 3 | US-09-151-011-3 |
| 5 | 397.6 | 23.7 | 1643 | 4 | US-09-343-623-3 |
| 6 | 397.6 | 23.7 | 1713 | 2 | US-08-486-839-5 |
| 7 | 397.6 | 23.7 | 1713 | 3 | US-09-151-011-5 |
| 8 | 397.6 | 23.7 | 1713 | 4 | US-09-343-623-5 |
| 9 | 397.2 | 23.6 | 1636 | 4 | US-09-039-198A-1 |
| 10 | 396 | 23.6 | 1656 | 4 | US-09-039-198A-3 |
| 11 | 384.6 | 22.9 | 1433 | 1 | US-08-694-915-1 |
| 12 | 360.6 | 21.5 | 1526 | 1 | US-08-694-915-3 |
| 13 | 73.6 | 4.4 | 147 | 4 | US-09-156-856-8 |
| 14 | 71.8 | 4.3 | 2452 | 2 | US-08-524-051-1 |
| 15 | 53.8 | 3.2 | 105 | 4 | US-09-156-856-13 |
| 16 | 48.4 | 2.9 | 108 | 4 | US-09-156-856-19 |
| 17 | 46.8 | 2.8 | 96 | 4 | US-09-156-856-17 |
| 18 | 44.2 | 2.6 | 3742 | 1 | US-08-694-915-5 |
| 19 | 44 | 2.6 | 1167 | 1 | US-07-939-501A-6 |
| 20 | 44 | 2.6 | 1167 | 4 | US-08-448-398-10 |
| 21 | 44 | 2.6 | 1320 | 1 | US-07-939-501A-14 |
| 22 | 44 | 2.6 | 1364 | 1 | US-07-939-501A-13 |
| 23 | 44 | 2.6 | 1405 | 1 | US-07-939-501A-9 |
| 24 | 44 | 2.6 | 1701 | 1 | US-07-939-501A-11 |
| 25 | 40.8 | 2.4 | 2712 | 3 | US-09-025-691-4 |
| 26 | 40 | 2.4 | 120 | 4 | US-09-156-856-18 |
| 27 | 40 | 2.4 | 132 | 4 | US-09-156-856-20 |

| | | | | | | |
|----|------|-----|-------|---|-------------------|-------------------|
| 28 | 38.8 | 2.3 | 1500 | 3 | US-09-052-778-1 | Sequence 1, Appl |
| 29 | 36.2 | 2.2 | 1074 | 1 | US-08-045-269C-3 | Sequence 3, Appl |
| 30 | 36.2 | 2.2 | 1074 | 3 | US-08-371-680-3 | Sequence 3, Appl |
| 31 | 36.2 | 2.2 | 1074 | 5 | PCT-US94-01198-3 | Sequence 3, Appl |
| 32 | 36.2 | 2.2 | 1554 | 1 | US-08-045-269C-1 | Sequence 1, Appl |
| 33 | 36.2 | 2.2 | 1554 | 3 | US-08-371-680-1 | Sequence 1, Appl |
| 34 | 36.2 | 2.2 | 1554 | 5 | PCT-US94-01198-1 | Sequence 1, Appl |
| 35 | 34.6 | 2.1 | 2934 | 3 | US-09-149-934-2 | Sequence 2, Appl |
| 36 | 34.4 | 2.0 | 354 | 2 | US-08-906-713-3 | Sequence 3, Appl |
| 37 | 34.4 | 2.0 | 1965 | 4 | US-09-178-252-26 | Sequence 26, Appl |
| 38 | 34 | 2.0 | 830 | 4 | US-08-998-416-557 | Sequence 557, App |
| 39 | 33.8 | 2.0 | 2831 | 2 | US-08-906-713-1 | Sequence 1, Appl |
| 40 | 33.4 | 2.0 | 31571 | 1 | US-08-323-443B-1 | Sequence 1, Appl |
| 41 | 33.4 | 2.0 | 53526 | 3 | US-08-658-136-2 | Sequence 2, Appl |
| 42 | 33.4 | 2.0 | 53526 | 3 | US-08-658-136-2 | Sequence 2, Appl |
| 43 | 33.4 | 2.0 | 53577 | 3 | US-08-658-136-1 | Sequence 1, Appl |
| 44 | 33.4 | 2.0 | 53577 | 3 | US-08-658-136-1 | Sequence 1, Appl |
| 45 | 33 | 2.0 | 4060 | 1 | US-08-308-949A-1 | Sequence 1, Appl |

ALIGNMENTS

RESULT 1

US-08-581-527-4

; Sequence 4, Application US/08581527

; Patent No. 5935798

; GENERAL INFORMATION:

; APPLICANT: Price, Paul A.

; APPLICANT: Johansen, Julia S.

; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER

; TITLE OF INVENTION: FOR DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson, P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: Windows95

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/581,527

; FILING DATE: 17-APR-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: WO95/01995

; FILING DATE: 19-JAN-1995

; APPLICATION NUMBER: 08/089,989

; FILING DATE: 09-JUL-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: Freeman, John W.

; REGISTRATION NUMBER: 29,066

; REFERENCE/DOCKET NUMBER: 07341/011001

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 617-543-5070

; TELEFAX: 617-542-8906

; INFORMATION FOR SEQ ID NO: 4:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1681 nucleic acids

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: genomic DNA

; IMMEDIATE SOURCE:

; CLONE: YKL-40

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 135....1681

Tue Dec 11 08:46:52 2001

US-08-581-527-4

Query Match 100.0%; Score 1681; DB 2; Length 1681;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ctaggtagtcgacacagagcggtgggcaaggagagccacacccctgcccctgctcg 60
DB 1 CTAGGTAGCTGGCACCAGAGCGGTGGGCAAGGAAGAGAGCCACACCCCTGCCCTGCTG 60

QY 61 ctgacgcaaatgggtgaaagcgctcacaacagcgcttgggtcctgggtgctgctcc 120
DB 61 CTGACGCCAGAAAGGGGTGGAAGCGCTCAACACAGGCTTGTGGTCCCTGGTGTCTCC 120

QY 121 agtgcgtctgcatacaaaactggtctgtactacacagcgtggtcccaagtgaggaag 180
DB 121 AGTGCTGCTGTCATACAAACTGGTGTCTACTTACACAGCTGGTCCCCAGTACCGGAAG 180

QY 181 gcatgggagcgtcttccagatgcccttgaccgcttccctgtgtacccacatcatata 240
DB 181 GCGATGGGAGCTGCTTCCAGATGCCCTTGACCGCTTCCTGTGTGTCACCATCATCTACA 240

QY 241 gcttgccaataaagcaacgatacacaacacacacacacacacacacacacacacac 300
DB 241 GCTTTGCCAATATAAGCAACAGATCACATGCACACCTGGGAGTGGATGATGTGACGCTCT 300

QY 301 acggcatgctcaacacactcaacacacacacacacacacacacacacacacacacac 360
DB 301 ACGGCATGCTCAACACACTCAACACACAGAACCCCAACCTGAAGACTCTCTGTGTCTCG 360

QY 420 gagatgaacttgggtctcaagatttccaaagatagcctcccaacacacacacacacac 420
DB 420 GAGATGAACATTTGGGTCTCAAGATTTCCAAAGATAGCTTCCAAACACACAGAGTGCCT 420

QY 480 ggaacttcaatcaagtacgcgcatcttctgcgcacacacacacacacacacacacacac 480
DB 480 GGACTTTTCATCAAGTCAGTACGGCATTTCTGCGCACCCCACTGGTGGTGGGTGAC 480

QY 540 ttgctgctctaccctgacgagagacacacacacacacacacacacacacacacacacac 540
DB 540 TTGCTGCTCTTACCCTGGAGGAGAGACAAACACCATTTTACCACCTAATCAAGGAAA 540

QY 600 tgaagcgcaattataaagaaagcccgagcgggaaagacagcctcgtcgcagcag 600
DB 600 TGAAGCGCAATTTATAAGGAAGCCAGCAGCGGAGGAAAAGCAGCTCCTGCTACGCGAG 600

QY 660 cactgtctgcggggaagtcaccattgacagcagctatgacattgccaaagatatcccaac 660
DB 601 CACTGTCTGCGGGGAAGGTCACATTTGACAGCAGCTATGACATTGCCAAGATATCCCAAC 660

QY 720 acctggatttcattagcatcatgacatgacatttccatggcgctggcggtgggacacag 720
DB 661 ACTGGATTTCATTAGCATATGACTAGCTAGCTATTTTCATGGCGCTGGCTGGGACACAG 720

QY 780 gccatcacagtcctccagcagcagcagcagcagcagcagcagcagcagcagcagcagcag 780
DB 721 GCCATCACAGTCCCCCTCAGCGGAGGTGAGGAGGATGCAAGTCTGTGACAGATTACGAACA 780

QY 840 ctgactatgctggtgggtacatgtgagcgtggggctcctccagtaagctgggtgag 840
DB 781 CTGACTATGCTGTGGGGTACATGTTGAGGCTGGGGGCTCCTGCCAGTAAGCTGGTGTGAG 840

QY 900 gcatccccacacctcgagagagattcactctgcttctctctctctctctctctctctctct 900
DB 841 GCATCCCCACCTTCGGGAGGAGCTTCACTGTGCTTCTTCTGAGACTGCTGTTCACAGCG 900

QY 960 caatctcaggaacgggaattccagccggttcccaagagcagcagcagcagcagcagcagcag 960
DB 901 CAATCTCAGGACCGGGAATTTCCAGCGCGGTTCACCAAGGAGGAGGAGCCCTTGCCTACT 960

QY 1020 atgagatctgtgacttctccgagacacagtcacatagaacccctcgccagcagcagtc 1020
DB 961 ATGAGATCTGTGACTTCTCTCCGCGAGCCACAGTCCATAGAACCTCGGCCAGCAGGTCC 1020

RESULT 2
PCT-US94-07754-4
; Sequence 4, Application PC/TUS9407754
; GENERAL INFORMATION:
; APPLICANT: THE REGENTS OF THE UNIVERSITY OF CALIFORNIA
; TITLE OF INVENTION: ASSAY FOR YKL-40 AS A MARKER FOR
; TITLE OF INVENTION: DEGRADATION OF MAMMALIAN CONNECTIVE TISSUE MATRICES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SPENSLEY HORN JUBAS & LUBITZ
; STREET: 1880 CENTURY PARK EAST, FIFTH FLOOR
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90067
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/07754

; FILING DATE: 08-JUL-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: HOWELLS, STACY L.
; REGISTRATION NUMBER: 34,842
; REFERENCE/DOCKET NUMBER: FD 3665
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619/455-5100
; TELEFAX: 619/455-5110
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1681 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; IMMEDIATE SOURCE:
; CLONE: YKL-40
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 135..1681
PCT-US94-07754-4

Query Match 99.6%; Score 1674.6; DB 5; Length 1681;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1677; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 ctagttagctggcaccagagcgctgggcaagggaagggcacaccctgcccctctctg 60
DB 1 CTAGGTAGCTGGCACAGGAGCGGTGGGCAAGGGAAGAGGCACACCCCTGCCCTGCTCTG 60
QY 61 ctgagcagaagaatgggtgaaggcggtctcaaaagggttggctgctggtgctgctcc 120
DB 61 CTGAGCGCAGAAATGGGTGTGAAGGCGTCTCAAAACAGGCTTTGTGTCTGTGGTGTCTCTCC 120
QY 121 aqctgctctctcatatacaactggctgctactacacagctgggtccagtagtcagggaag 180
DB 121 AGTGCTGCTCTGCATACAAACTGTGCTGTACTACACACAGCTGGTCCCNAGTACCGGGAAG 180
QY 181 gcatggagagctgtctccagatgcccctgaccgcttctctgctgtgtgtaccacacatctaca 240
DB 181 GCGATGGGAGCTGCTTCCAGATGCCCTTGACCGCTTCTGTGTACCCACATCATCTACA 240
QY 241 gcttggccaataaagcaacgcatcacaacacacacacacacacacacacacacacacacacac 300
DB 241 GCTTTGCCAATAAAGCAACGATCATCGACACCTGGGAGTGGAAATGATGTGACGCTCT 300
QY 301 acgagcatctcaacacactcaacaacacagaaaccccaacacacacacacacacacacacacac 360
DB 301 ACGGCATGCTCAACACACTCAACAACAGAACCCCAACCTGAAGACTCTCTGTCTGTCTG 360
QY 361 gaggatgaaacttgggtctcaagaattttccaagatagcctcccaacacacacacacacacacac 420
DB 361 GAGGATGGAACCTTTGGGTCTCAAGATTTTCCAAGATAGCTCCCAACACCCAGAGTCGCC 420
QY 421 ggaatttcaataagtcagtagccgcatcttctgcacccaatggcttggatggggtgacc 480
DB 421 GGACTTTTCATCAAGTCAGTACCGCCATTTCTCGCACCCCATGGCTTTGTATGGGCGGTGACC 480
QY 481 ttgcttggtctctaccctggagcgagagacaacacacacacacacacacacacacacacacacac 540
DB 481 TTGCTTGCTCTPACCTTGACCGGAGAGACAAACACCATTTTACCACCTPAATCAAGGAAA 540
QY 541 tgaaggccgaatttataagggaagcccgccaggggaagaaagcagctcctgctcagcgag 600
DB 541 TGAAGGCCGAATTTATAAGGAAGCCAGCCAGGGGAAAAGACGCTCCTGCTCAGCGCAG 600
QY 601 cactgtctcgggggaaggttcacattgacagcagctatgacatttgccaagatatcccaac 660
DB 601 CACTGTCTCGGGGAAGGTTCACATTGACAGCAGCTATGACATTGCAAGATATCCCAAC 660
QY 661 acctggatttcattagcatcatgacctaagattttcatggcgctggcggtgggacacag 720

DB 661 ACCTGGATTTTATTAGCATCATGACTACGATTTTCATGGCCCTGGCGTGGGACACAG 720
QY 721 gccatcacagtcctcctcaaggcgaggtggaagatgcaagtcctcacagattcacaaca 780
DB 721 GCCATCACAGTCCCTCTGTCCGAGGTGAGGAGTGCAGAGTCTCTGACAGATTCAGACA 780
QY 781 ctgactatgctgtgggttacatgttgaggctggggctcctccagtagtaagctgggtgatgg 840
DB 781 CTGACTATGCTGTGGGTACATGTTGAGGCTGGGGCTCCTGCCAGTAGCTGGTGATGG 840
QY 841 gcattccacaccttcgggaggagcttcaactctggtcttctctgagactggtgttccagggc 900
DB 841 GCATCCCCACCTTCGGGAGGAGCTTCACTGTGCTCTTCTGAGACTGGTGTTCAGCGC 900
QY 901 caatctcaggaccgggaatttcaggccgggttcaccaaggaggcaggacccttgcctact 960
DB 901 CAATCTCAGGACCGGGAATTCAGGCCGGTTCCACCAAGGAGGACGCTTGGCTACT 960
QY 961 atgagatctgtgacttctcccgcgagccacagtcataagaacctcgccagcaggtcc 1020
DB 961 ATGAGATCTGTGACTTCTCCCGGAGGCCACAGTCCATAGAACCTTCGCCAGCAGGTCC 1020
QY 1021 cctatgccaccaaggggcaaccagtggttaggatacgaagaccaggaagcgtcaaaagca 1080
DB 1021 CCTATGCCACCAAGGCAACCACTGGGTAGGATACGACGACGAGAAAGCGTCAAAGCA 1080
QY 1081 aggtcagtagtaccatgaagtagtagcagctggcaggcgccatgtagtggccctggacctgg 1140
DB 1081 AGGTGACGTACCTGAAGGATAGGACGTGGCAGGCGCCATGATATGGGCCCTGGACCTGG 1140
QY 1141 atgacttccagggtcctcttcgcccaggtatctgcgtcctcctcaccatgacctca 1200
DB 1141 ATGACTTCCAGGGCTCTCTCTCGGCCAGGATCTGGCTTCCCTCTCACCATATGCCATCA 1200
QY 1201 aggatgactcgtctcaactgtagcctctgtctgcacacagcagcgggggccaagatgc 1260
DB 1201 AGGATGCACTCGCTGCAAGCTAGCCCTCTGTCTGTCACACACGACGCGGGGCCAAGGATGC 1260
QY 1261 cccgtccctcgtgctggcggcgagcctgatacactgcccctgctgagtcaccaggtgag 1320
DB 1261 CCGCTCCCGCTGTGCTGGCTGGCGGAGCCTGATCACCTGCCCTGCTGAGTCCCGGCTGAG 1320
QY 1321 cctcagttcctcctccttggggcctatgcagaggtcccaacacacacacagattgagctcag 1380
DB 1321 CCTCAGTCTCCTCTCCCTTGGGGCCTATGACAGAGGTCCCAACACACAGATTTGAGCTCAG 1380
QY 1381 ccttggtggcgagaggttacacacttgttgatgatttaaggaaaagtgttacagatcccc 1440
DB 1381 CCTTGGTGGCAGAGAGGTACACACTTGTGTGATGATTAATGGAAATGTTTACAGATCCCC 1440
QY 1441 aagcctggcagggaatttcttcaactcctgcctcctcctcctcctcctcctcctcctcctcctc 1500
DB 1441 AAGCCTGGCAAGGAATTTCTTCACTCCCTGCCCTAGCCCTCTCTTATCAAAAGACAC 1500
QY 1501 cattttgcaagctctatcaccaaggagcacaacatctcacaagacacacagtagaccatact 1560
DB 1501 CATTTGGCAAGCTCTATCACCAGGAGGCAACATCTTACAGACACAGTAGTACCATACT 1560
QY 1561 aattatcccccctgcaaacgagccttgaaccttcaacttaggaacgtaactcgtgtccccct 1620
DB 1561 AATTATACCCCTGCAAAAGCCAGCTTGAACCTTCACTTAGGAACGTAATCGTGTCCCT 1620
QY 1621 atcctacttcccccttctcaattccacagctgctcctaataaagataaagagtttaacagtg 1680
DB 1621 ATCTACTTCCCTCTCTTAATTCACAGCTGCTCAATAAAGTACAAGAGTTTAAACAGTGT 1680
QY 1681 g 1681
DB 1681 G 1681

US-09-343-623-3
HYPOTHETICAL: NO

Query Match 23.78; Score 397.6; DB 4; Length 1643;
 Best Local Similarity 62.4%; Pred. No. 7.2e-104;
 Matches 701; Conservative 0; Mismatches 399; Indels 24; Gaps 4;

| | | | |
|----|------|----------------------------------------------------------------|------|
| QY | 68 | cagaatgggtgtgaagcgctgtcaaacagagcttctgtgtcctgtgtgtcctccagtgctg | 127 |
| DB | 9 | CATCATGGTGGGTCTGTGGCTGGGACAGTTTCATGTCTCCTGTGATGATCCATGGGG | 68 |
| QY | 128 | ctctgcatacaaatgttgtctactacacacagctgtgtccagtagtccaggggaagcgatgg | 187 |
| DB | 69 | CTCTGCTCCAAACTGGTCTGTACTTCACCAACTGGGCCCGGAGTACAGACAGGGGAGGC | 128 |
| QY | 188 | gagctgtctccagatgccttgaccgcttctctgtgtaccacatcatctacagctttgc | 247 |
| DB | 129 | TGCGTCTGTGCCAAGGACTTGGACCCAGCCCTTGTGACCCACCTCATCTACGCCTCG | 188 |
| QY | 248 | caataaagcaacgatcacatcgacacctggagtggaatgatgtgaogctctacggcat | 307 |
| DB | 189 | TGGCATGACCAACCACAGCTGAGCACCACCTGAGTGGATGACGAGACTCTCTACAGGA | 248 |
| QY | 308 | gtcaacacacatcaaacacacacacacacacacacacacacacacacacacacacacac | 367 |
| DB | 249 | GTTCAATGGCCTGAAGAAGATGAATCCCAAGCTGAAGACCCCTGTAGCCATCGGAGGCTG | 308 |
| QY | 368 | gaacttgggtctaaagtatttccaagtagcctcaacacacacacacacacacacacacac | 427 |
| DB | 309 | GAATTTCGGCATCTAGAGATTTCACAGATATGTAGCCACGCGCCAAACCGTCAGACCT | 368 |
| QY | 428 | catcaagtcaagtcacgccttctctgcgaccccatggcttggatggcggtgacactgcctg | 487 |
| DB | 369 | TGTCAACTCGGCCATCAGGTTCTTGGCGAAATACAGACTTGTACGGCTTGTACCTTGACT | 428 |
| QY | 488 | gctctacctgtgagc-----gagagcaaacacacatttaccaccccta | 532 |
| DB | 429 | GGAGTACCCAGGAAGCAGGGGAGCCCTGCCGTAGACAAGAGCGCTTCAACACCTTGT | 488 |
| QY | 533 | caaggaatgaagggcgaatttaaaaggaagccag---ccagggaaaaagcagctcct | 589 |
| DB | 489 | ACAGGACTTGGCCAAATGCCCTTCAGCAGGAAGCCACAGACCTCAGGGAAGAACGCCCTCT | 548 |
| QY | 590 | gtcagcgcagcactgtctgctgggggaaggtccaccatgcagcagcgtatgacattgcaaa | 649 |
| DB | 549 | TCTGAGTGCAGCGGTCCAGCTGGCAGACCTATGTGATGCTGGATACGAGGTGGACAA | 608 |
| QY | 650 | gatataccacacactggatttcattagcatcatgacctagctatttcatggcgctggcg | 709 |
| DB | 609 | AATCGGCCAAGCACTGGATTTGTCAACCTTATGGCCCTACACTTCATGGCTCTTGGGA | 668 |
| QY | 710 | tgggaccagagccatcacactgcctccagcgaggttcaggaggatgcgaagtctgcag | 769 |
| DB | 669 | GAAAGTCAAGGGACATTAACAGCCCTCTACAGAGGCAAGAGAGAGTGTGCAGCAGC | 728 |
| QY | 770 | attcagcaacactgactatgtgtgtgggttacatgttgagctgggggctcctgcagtaa | 829 |
| DB | 729 | CAGCCTCAACGTGGATGCTGTGTGCAACAGTGGCTGCAGAAAGGGGACCCCTGCCAGCA | 788 |
| QY | 830 | gctgtgtgatgggcatcccaactctggagagagcttcaacttg---cttctctgagac | 886 |
| DB | 789 | GCTGATCCTTGGCATGCGCTACCTACGGACAGCTCTCTTCACATGGCCCTCTCATCAGAC | 848 |
| QY | 887 | tgtgtgtccagcgcaatctcaggagaccgggaattccaggcggttccaaagaggagcg | 946 |
| DB | 849 | CAGAGTGGGGCCCCAGCCACAGGGTCTGGCACCTCCAGGCCCTTCCAAAGGAAGGAGG | 908 |
| QY | 947 | gacctgtcctactatgatgtctgtgaacttctccgcggagccacagtcctataaacct | 1006 |
| DB | 909 | GATGCTGCGCCTACTATGAAGTCTG---CTCTGGAAAGGGGGGCCCAACACAGAAATCA | 965 |
| QY | 1007 | cgccagcagctgccctatgcaccaagggcaaacagtggtgtaggtatgcacgaccagga | 1066 |

489 ACAGGACTTGGCCATGCTTCCACAGGAAGCCAGACCTCAGGAAAGGACGCTTCT 548
QY 590 gctcagcagcagctctgcggggaagtgtaacattgacagcagctatgacattggccaa 649
Db 549 TCTGAGTGCAGCGGTTCACGCTGGGAGACCTATGTGATGCTGGATACGAGGTGGACAA 608
QY 650 gatatacaaacactgatttattagatcatcatgacactacgtatcttgcgctgctg 709
Db 609 AATCGCCAGAACCTGGATTGTCAACCTTATGGCTACGACTTCCATGGCTCTGGGA 668
QY 710 tgggacacagcagccatcaacagtcctccctcggcgaggtcagggagatgcaagtcctgacag 769
Db 669 GAAGGTCACGGGACATAAAGCCCTCTACAGAGGCAAGAGAGAGTGGTGCAGACAC 728
QY 770 attcagcaaacactgactatgctgtgggtacatgttgaggctgggggctcctgcagtaa 829
Db 729 CAGCCTCAACGTGGATGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAA 788
QY 830 gctgggtgagggcatcccccaccttcgggagagcttcaactctgg----cttctctgagac 886
Db 789 GCTGATCTTGGCATGCTACCTACGAGACGCTCTTACACACTGGCCTCCTCATCAGACAC 848
QY 887 tgggttccagcgcacatcagcagcgggaattccagccggttccacaaagagcag 946
Db 849 CAGAGTGGGGGCCCCAGCCACAGGGGTCTGGCACTCCAGGCCCTTACCAGGAAGAGG 908
QY 947 gacccctgctactatgactctgacttctcctccgagagccacagctccatagaacct 1006
Db 909 GATGCTGCTTACCTATGAAGTCTG---CTCTGGAAGGGGCCACCAACAGAGATCCA 965
QY 1007 cggcagcagctccctatgcccacaaagggcagcagctgggttaggatacagcagcagga 1066
Db 966 GGATCAGAAAGTGGCTTACATCTTCCGGGACCAACAGTGGGTGGCTTTGATGATGGA 1025
QY 1067 aagcgttaaaagcaggtgcagctaccgaaagatgagcagctgcagcgcagcagctggtatg 1126
Db 1026 GAGCTTCAAAACCAAGGTCAGCTATCTGAAGCAGAAAGGACTGGCGGGGCCATGGTCTG 1085
QY 1127 ggcctgacactgactcctccagggctcctctctgcggccag 1170
Db 1086 GGCACCTGACTAGATGACTTGGCGGCTTCTCCTGCAACCAAG 1129

RESULT 8
US-09-343-623-5
; Sequence 5, Application US/09343623
; Patent No. 6303118
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: A human chitinase, its recombinant
; TITLE OF INVENTION: production, its use for decomposing chitin, its use
; TITLE OF INVENTION: in therapy or prophylaxis against infection diseases.
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hofmann & Baron
; STREET: 350 Jericho Turnpike
; CITY: Jericho
; STATE: New York
; COUNTRY: United States of America
; ZIP: 11758
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; ATTORNEY/AGENT INFORMATION:
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/343,623
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/486,839
; FILING DATE: 07-June-1995

NAME: Baron, Ronald J.
REGISTRATION NUMBER: 29,281
REFERENCE/DOCKET NUMBER: 294-26
TELECOMMUNICATION INFORMATION:
TELEPHONE: (516) 822-3550
TELEFAX: (516) 822-3582
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1713 base pairs
TYPE: nucleic acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: cdna
HYPOTHETICAL: NO
US-09-343-623-5
Query Match 23.7%; Score 397.6; DB 4; Length 1713;
Best Local Similarity 62.4%; Pred. No. 7.4e-104;
Matches 701; Conservative 0; Mismatches 399; Indels 24; Gaps 4;
QY 68 cagaatgggtgtgaagcgctctcaaacaggtttgtggtcctggtgctgctccagtgctg 127
Db 9 CATCATGTGTCGCTGCTGTGGCCTGGGCAAGGTTTCATGGTCTTGTGATGATCCCATGGGG 68
QY 128 ctctgcatacaaacactgctgtctactacacacagctggtcccgatgacccgggaaggcgatgg 187
Db 69 CTCTGCTCCAAACTGCTGTCTACTTACCAACTGGGCCAGTACAGACAGAGGGGAGGC 128
QY 188 gactgtcttccagatgacctgacccgttctctgtgtaccacacatcatcagctttgc 247
Db 129 TCGCTTCTTCCCAAGGACTTGGACCCCAAGCCTTTCACCCACCTCATCTACGCCCTTCGC 188
QY 248 caataaagaacagatcacatcgacacctgggagtggaatgagtgtgacgtctctacggcat 307
Db 189 TGGCATGACCAACACACAGCTGAGCACCTGAGTGGTGAATGACGAGACTCTCTFACCAGGA 248
QY 308 gctcaacacactcaacacacacacacacacacacacacacacacacacacacacacacac 367
Db 249 GTTCATGGCTTGAAGAAGATGAATCCCAAGCTGAAGACCTGTTAGCCCATCGGAGGCTG 308
QY 368 gaacttgggtgtctaaagattttccaaagatgagctcctccttctgtcgcgagatg 427
Db 309 GAATTCGGCACTCAGAAGTTCAGATATGGTAGCCGCGCCCAACACCCGTGAGACCTT 368
QY 428 catcaagtacgtaccgaccttctgcacacacacacacacacacacacacacacacacacac 487
Db 369 TGCAACTCGGCCCATCAGGTTTCTGCGCAATACAGCTTTGACGGCCTTGACCTTGACTG 428
QY 488 gctctacccctggagc-----gagagacacacacacacacacacacacacacacacac 532
Db 429 GGAGTACCCAGGAAGCCAGGGGAGCCCTGCGGTAGACAAGGAGCGCTTCACACCCCTGGT 488
QY 533 caaggaatgaagccgcaattataaaggaagccag---ccagggaaaaagcagctcct 589
Db 489 ACAGGACTTGGCCCAATGCCCTTCCAGAGGAAGCCCAAGACCTCAGGGGAAGGACCCCTCT 548
QY 590 gctcagcagcagcagctgtcgcggggaaggtccaccatgacagcagcagctatgacattgcaa 649
Db 549 TCTGAGTGCAGCGGTTCAGAGTGGGACAGCTATGTGGATGCTGGATACGAGGTGGACAA 608
QY 650 gatatacaaacactgagtttattagatcatcatgacactacatttcatgctgctgctg 709
Db 609 AATCGCCAGAACCTTGGATTTGTCAACCTTATGGCTTACGACTTCCATGGCTCTTGGGA 668
QY 710 tgggacacagcagccatcagctccctcagcgaggtcagggagatgcaagtcctgacag 769
Db 669 GAAGGTCACGGGACATAAAGCCCTCTACAGAGGCAAGAGAGAGTGGTGCAGACAGC 728
QY 770 attcagcaaacactgactatgctgtgggtacatgttgaggctgggggctcctgcagtaa 829
Db 729 CAGCCTCAACGTGGATGCTGTGCAACAGTGGCTGCAGAGGGGACCCCTGCCAGCAA 788

| | | | |
|----|------|-----------------------------------------------------------------------|------|
| Qy | 830 | gctgtgtgtagtggggaatcccccaactcttcggagagagcttcaactctgg---cttcttcttgagac | 886 |
| | | | |
| Db | 789 | gcctgactcttgccatgccttaccctacggagcgcctcttccacattgcctcctcatcagacac | 848 |
| | | | |
| Qy | 887 | tgtgtgttcagagcccaatctcaagcaccgggaattccagccgggtttcaccgaaggagccagg | 946 |
| | | | |
| Db | 849 | cagagctggggggccccagccacagggtcttggcactccaggccctttccacgaaggaaggagg | 908 |
| | | | |
| Qy | 947 | gaccccttgctactatgagatctgtgactctctccgcggagccacagtcacatagaacccct | 1006 |
| | | | |
| Db | 909 | gatgtctggccctactatgaagtctg---ctctctggaaaggccgccacacacagagaaatcca | 965 |
| | | | |
| Qy | 1007 | cggccagcaggtccctctatgccaccaaggcaaccagctgggtaggatctacgacgaccagga | 1066 |
| | | | |
| Db | 966 | ggatcagaaggctgcccttacctctccgggacacacacagctgggtggcctttgatgatgtgga | 1025 |
| | | | |
| Qy | 1067 | aagcgtcaaaagcaagtgtagtacctgaagataggcagctggcagcgccactggatggtatg | 1126 |
| | | | |
| Db | 1026 | gagcttcaaaaccaggctcagctatctgaagcagaagggactggcgggggccactgggtctg | 1085 |
| | | | |
| Qy | 1127 | ggccctggacctggatgactctcagggtctctctctgcggccagg | 1170 |
| | | | |
| Db | 1086 | ggcactctggactttagatgacttttgcgggctttctcttcgaacacagg | 1129 |
| | | | |

RESULT 9

```

US-09-039-198A-1
: Sequence 1, Application US/09039198A
: Patent No. 6200951
: GENERAL INFORMATION:
: APPLICANT: Gray, Patrick W.
: APPLICANT: Tjolsker, Larry W.
: TITLE OF INVENTION: CHITINASE CHITIN-BINDING FRAGMENTS
: NUMBER OF SEQUENCES: 34
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 233 South Wacker Drive/6300 Sears Tower
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/039,198A
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Rin-Laures, Li-Hsien
: REGISTRATION NUMBER: 33,547
: REFERENCE/DOCKET NUMBER: 27866/34391
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (312) 474-6300
: TELEFAX: (312) 474-0448
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1636 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 2..1399
: FEATURE:
: NAME/KEY: mat_peptide
: LOCATION: 65..1399
: US-09-039-198A-1

```


| | | | |
|----|-----|----------------------------------------------------------------|-----|
| QY | 127 | gctctgcatacaaaactgctgtctactacacacagctggtccacgtacacgggaagcgatg | 186 |
| DB | 200 | GATCTGCCTACAAACTGGTTTGCTACTTTTACCAACTGGTCCGACGACGGGAGGAACGAG | 259 |
| QY | 187 | ggagctgcttccacagatgccccttgaccgcttctctgtacccacatcatctacagctttg | 246 |
| DB | 260 | GAAAAATTACCCCTGAGAATATTGACCCCTTCCCTATGCTCTCATCTCATCTATTCATTCG | 319 |
| QY | 247 | ccaataaagcaacagatcacatcgacacctggagtggaatgagtgcacgtctctacggca | 306 |
| DB | 320 | CCAGCATCGAAACAACAGGTTTATCATCAAGGACACAGAGTGAAGTGATGCTTACCAGA | 373 |
| QY | 307 | tgctcaacacactcaacacacgaaccccaacctgaagactctgtctctctcgagatg | 366 |
| DB | 380 | CCATCAACAGCTCTCAAAACCAAGAATCCCAAACTGAAAAATTCCTGTGCCATTGGAGGGT | 439 |
| QY | 367 | ggaacttgggtctcaaaagatttccaaagatagcctccacacccagagtcgcgcgaact | 426 |
| DB | 440 | ACCTGTTTGGTTCCAAAGGTTCCACCCCTATGGTGGATTCCTTACATCACGCTTGGAAAT | 499 |
| QY | 427 | tcatacagtcagtcaccgccattctctgcgacccatgctgttgatgggcgtgaccttgct | 486 |
| DB | 500 | TCATTAACTCCATAATCCTGTTCTTGAGGAACCATAACTTGTATGACATGGAGTGTAAAGCT | 559 |
| QY | 487 | ggctctacctggcagcgagagacacacacacattttaccaccttaatacagggaataagag | 546 |
| DB | 560 | GGATCTACCCAGATCGAAGAGAAACACTCATTTTCACTGTGTGATTCATGATGTAGCAG | 619 |
| QY | 547 | ccgaatttataaaggaagcccgacggaggaaaaagca---gctcgtctcgcgcgacgac | 603 |
| DB | 620 | AAGCCTTTCAGAGGACTTCACAAAATCCACCAAGGAAGGCTTCTTGTAGCTGGGGGG | 679 |
| QY | 604 | tgctgcgggaaggtcaccaattgacagagcgtatgacattgccaaagatatcccaaccc | 663 |
| DB | 680 | TATCTGCAAGGAGGCAATGATTGATAACAGCTATCAAGTTGTAGAACACTGGCAAAAGATC | 739 |
| QY | 664 | tgagattcttagcatcatgacctacgatttctcatgcgcctg-----gcgtgggagcca | 717 |
| DB | 740 | TGGATTTTCATCAACTCCTGTGCTTTGACITTCATGGTCTTGGGAAAAAGCCCCCTTATCA | 799 |
| QY | 718 | caggccatcacgtccctccaggcgaggtcagggagatgcaagtccctgcagattcagca | 777 |

Tue Dec 11 08:46:52 2001

us-09-164-862b-4.rni

```
QY 838 tgggcatccacccttcgggagagagcttcaactcgtgcttct-----tcigagactggtgtt 893
Db 827 TCGGTATTCCTACGTCGATCTTACCCCTATCTGCTGGCAACAACAACACTACGGTC 886
QY 894 ccagcccaatctcaggaccg-----ggaattccaggccgggtccaccaaggagg 942
Db 887 TCGGCACCTTCATCACAGGAGCAGGCGGGGTGACCCCTGGCCCATACACCAATGCTA 946
QY 943 caggacccttgctactatgatgtgtgacttct-----ccgcgaggacca 990
Db 947 CAGGATTTGGGCTTATTATGAGATCTGACAGAGTAGACAAGGATGACTCCGGCTGGA 1006
QY 991 cagtccatagaaccctcggccagcaggtccctctatcccaaggcgaaccagtggttag 1050
Db 1007 CGAAGAAATGGGACGAGCAAGGCAAGTGCCCTTATCCCTACAGGGCACCAGTGGGTTG 1066
QY 1051 gatacagcagcaggaaagcgtcaaaagcaaggtgcagtacctgaaggataggcagctgg 1110
Db 1067 GATACGAGAACCTTCGAGCGGTGGAGATCAAGATGAAC TGGATTAAACAGAGGATACC 1126
QY 1111 caggcgcctatggtatggccctggacctgatgacttccagggtcctcttcgcggccagg 1170
Db 1127 TTGGAGCCATGACTTGGGCTATCGACATGATGACTTCCCAAGGACTGTGTGGAGAGAAGA 1186
QY 1171 atctgcgttccctctcaccatgccatcaaggatgactcgtcgtgcaacgtagccctctg 1230
Db 1187 ACCAATTGATCAAGATTCTTCATAAGCACATGAGCTTTACACAGTCCCGCCTCCTCATA 1246
QY 1231 ttctgcacagcagcggggccgaaggatgcccgcgtcccc 1269
Db 1247 CAGAGACACACACCGACTCTCTGAATGGGCCCGTCCAC 1285
```

```
RESULT 15
US-09-156-856-13
; Sequence 13, Application US/09156856A
; Patent No. 6221591
; GENERAL INFORMATION:
; APPLICANT: Aerts, Johannes M.
; TITLE OF INVENTION: Determination of a genetic risk factor for infection
; TITLE OF INVENTION: and other diseases, and detection of activated
; TITLE OF INVENTION: phagocytes
; FILE REFERENCE: Sequence 1-20
; Patent No. 6221591
; CURRENT APPLICATION NUMBER: US/09/156,856A
; CURRENT FILING DATE: 1998-09-18
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-156-856-13
```

```
Query Match      3.2%; Score 53.8; DB 4; Length 105;
Best Local Similarity 72.2%; Pred. No. 2.7e-06;
Matches 70; Conservative 0; Mismatches 27; Indels 0; Gaps 0;
```

```
QY 1074 aaagcaaggtgcagctacgtgaagatagcagctggcaggcccatggtatgggcccctg 1133
Db 1 aaacccaaggtcagctatctctgaagcagaaggactggcggggggcccatggtctgggcaactg 60
QY 1134 gacctggtgacttccagggtcctctctcgcggccagg 1170
Db 61 gacttagatgacttgcgggtctctcctcgtcaaccagg 97
```

Search completed: December 6, 2001, 23:38:35
Job time: 4420 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2001, 22:32:00 ; Search time 2421.7 seconds
(without alignments)
12165.785 Million cell updates/sec

Title: US-09-164-862B-4
Perfect score: 1681
Sequence: 1 ctaggtagctgcaccagga.....tacaagagttaacagtg 1681

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 17159718 seqs, 8763200856 residues
Total number of hits satisfying chosen parameters: 34319436

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

| | |
|-----|-----------------------------------------|
| 1: | /cgn2_6/ptodata/1/pna/PTCUS_COMB.seq:* |
| 2: | /cgn2_6/ptodata/1/pna/US06_COMB.seq:* |
| 3: | /cgn2_6/ptodata/1/pna/US07_COMB.seq:* |
| 4: | /cgn2_6/ptodata/1/pna/US080_COMB.seq:* |
| 5: | /cgn2_6/ptodata/1/pna/US081_COMB.seq:* |
| 6: | /cgn2_6/ptodata/1/pna/US082_COMB.seq:* |
| 7: | /cgn2_6/ptodata/1/pna/US083_COMB.seq:* |
| 8: | /cgn2_6/ptodata/1/pna/US084_COMB.seq:* |
| 9: | /cgn2_6/ptodata/1/pna/US085_COMB.seq:* |
| 10: | /cgn2_6/ptodata/1/pna/US086_COMB.seq:* |
| 11: | /cgn2_6/ptodata/1/pna/US087_COMB.seq:* |
| 12: | /cgn2_6/ptodata/1/pna/US088_COMB.seq:* |
| 13: | /cgn2_6/ptodata/1/pna/US089_COMB.seq:* |
| 14: | /cgn2_6/ptodata/1/pna/US090_COMB.seq:* |
| 15: | /cgn2_6/ptodata/1/pna/US091_COMB.seq:* |
| 16: | /cgn2_6/ptodata/1/pna/US092_COMB.seq:* |
| 17: | /cgn2_6/ptodata/1/pna/US093_COMB.seq:* |
| 18: | /cgn2_6/ptodata/1/pna/US094_COMB.seq:* |
| 19: | /cgn2_6/ptodata/1/pna/US095A_COMB.seq:* |
| 20: | /cgn2_6/ptodata/1/pna/US095B_COMB.seq:* |
| 21: | /cgn2_6/ptodata/1/pna/US095C_COMB.seq:* |
| 22: | /cgn2_6/ptodata/1/pna/US095D_COMB.seq:* |
| 23: | /cgn2_6/ptodata/1/pna/US096A_COMB.seq:* |
| 24: | /cgn2_6/ptodata/1/pna/US096B_COMB.seq:* |
| 25: | /cgn2_6/ptodata/1/pna/US096C_COMB.seq:* |
| 26: | /cgn2_6/ptodata/1/pna/US096D_COMB.seq:* |
| 27: | /cgn2_6/ptodata/1/pna/US096E_COMB.seq:* |
| 28: | /cgn2_6/ptodata/1/pna/US097A_COMB.seq:* |
| 29: | /cgn2_6/ptodata/1/pna/US097B_COMB.seq:* |
| 30: | /cgn2_6/ptodata/1/pna/US097C_COMB.seq:* |
| 31: | /cgn2_6/ptodata/1/pna/US098_COMB.seq:* |
| 32: | /cgn2_6/ptodata/1/pna/US099_COMB.seq:* |
| 33: | /cgn2_6/ptodata/1/pna/US6000_COMB.seq:* |
| 34: | /cgn2_6/ptodata/1/pna/US6001_COMB.seq:* |
| 35: | /cgn2_6/ptodata/1/pna/US6002_COMB.seq:* |
| 36: | /cgn2_6/ptodata/1/pna/US6003_COMB.seq:* |
| 37: | /cgn2_6/ptodata/1/pna/US6004_COMB.seq:* |
| 38: | /cgn2_6/ptodata/1/pna/US6005_COMB.seq:* |
| 39: | /cgn2_6/ptodata/1/pna/US6006_COMB.seq:* |
| 40: | /cgn2_6/ptodata/1/pna/US6007_COMB.seq:* |
| 41: | /cgn2_6/ptodata/1/pna/US6008_COMB.seq:* |
| 42: | /cgn2_6/ptodata/1/pna/US6009_COMB.seq:* |
| 43: | /cgn2_6/ptodata/1/pna/US6010_COMB.seq:* |

| | |
|-----|-----------------------------------------|
| 44: | /cgn2_6/ptodata/1/pna/US6011_COMB.seq:* |
| 45: | /cgn2_6/ptodata/1/pna/US6012_COMB.seq:* |
| 46: | /cgn2_6/ptodata/1/pna/US6013_COMB.seq:* |
| 47: | /cgn2_6/ptodata/1/pna/US6014_COMB.seq:* |
| 48: | /cgn2_6/ptodata/1/pna/US6015_COMB.seq:* |
| 49: | /cgn2_6/ptodata/1/pna/US6016_COMB.seq:* |
| 50: | /cgn2_6/ptodata/1/pna/US6017_COMB.seq:* |
| 51: | /cgn2_6/ptodata/1/pna/US6018_COMB.seq:* |
| 52: | /cgn2_6/ptodata/1/pna/US6019_COMB.seq:* |
| 53: | /cgn2_6/ptodata/1/pna/US6020_COMB.seq:* |
| 54: | /cgn2_6/ptodata/1/pna/US6021_COMB.seq:* |
| 55: | /cgn2_6/ptodata/1/pna/US6022_COMB.seq:* |
| 56: | /cgn2_6/ptodata/1/pna/US6023_COMB.seq:* |
| 57: | /cgn2_6/ptodata/1/pna/US6024_COMB.seq:* |
| 58: | /cgn2_6/ptodata/1/pna/US6025_COMB.seq:* |
| 59: | /cgn2_6/ptodata/1/pna/US6026_COMB.seq:* |
| 60: | /cgn2_6/ptodata/1/pna/US6027_COMB.seq:* |
| 61: | /cgn2_6/ptodata/1/pna/US6028_COMB.seq:* |
| 62: | /cgn2_6/ptodata/1/pna/US6029_COMB.seq:* |
| 63: | /cgn2_6/ptodata/1/pna/US6030_COMB.seq:* |
| 64: | /cgn2_6/ptodata/1/pna/US6031_COMB.seq:* |
| 65: | /cgn2_6/ptodata/1/pna/US6032_COMB.seq:* |

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match % | Length | IDB | ID | Description |
|------------|--------|---------------|--------|-----|---------------------|-------------------|
| 1 | 1681 | 100.0 | 1681 | 15 | US-09-164-862B-4 | Sequence 4, Appli |
| 2 | 1681 | 100.0 | 1681 | 16 | US-09-215-077A-4 | Sequence 4, Appli |
| 3 | 1575.4 | 93.7 | 1741 | 14 | US-09-023-655-1353 | Sequence 1353, Ap |
| 4 | 1575.4 | 93.7 | 1925 | 58 | US-60-257-537-104 | Sequence 104, App |
| 5 | 1575.4 | 93.7 | 1925 | 58 | US-60-257-537-115 | Sequence 115, App |
| 6 | 1561.2 | 92.9 | 1797 | 1 | PCT-US01-14827-6232 | Sequence 6232, Ap |
| 7 | 1561.2 | 92.9 | 1797 | 22 | US-09-577-408-6988 | Sequence 6988, Ap |
| 8 | 1548.2 | 92.1 | 1964 | 30 | US-09-770-173-3081 | Sequence 3081, Ap |
| 9 | 1548.2 | 92.1 | 1964 | 30 | US-09-770-174-4191 | Sequence 4191, Ap |
| 10 | 1542 | 91.7 | 1739 | 49 | US-60-164-285-3401 | Sequence 3401, Ap |
| 11 | 1540.6 | 91.6 | 1547 | 1 | PCT-US01-14827-6231 | Sequence 6231, Ap |
| 12 | 1540.6 | 91.6 | 1547 | 22 | US-09-577-408-6987 | Sequence 6987, Ap |
| 13 | 1536.8 | 91.4 | 1786 | 17 | US-09-359-922-6683 | Sequence 6683, Ap |
| 14 | 1536.8 | 91.4 | 1786 | 17 | US-09-359-922-6683 | Sequence 6683, Ap |
| 15 | 1521.2 | 90.5 | 1824 | 60 | US-60-278-258-10870 | Sequence 10870, A |
| 16 | 1469.8 | 87.4 | 1825 | 50 | US-60-172-373-8932 | Sequence 8932, Ap |
| 17 | 1425.2 | 84.8 | 2055 | 54 | US-60-213-360-5288 | Sequence 5288, Ap |
| 18 | 1290.6 | 76.8 | 1610 | 30 | US-09-760-475-882 | Sequence 882, App |
| 19 | 1128 | 67.1 | 1152 | 58 | US-60-257-537-105 | Sequence 105, App |
| 20 | 1128 | 67.1 | 1152 | 58 | US-60-257-537-106 | Sequence 106, App |
| 21 | 1058 | 62.9 | 1355 | 58 | US-60-257-537-116 | Sequence 116, App |
| 22 | 1035 | 61.6 | 1072 | 60 | US-60-278-258-10878 | Sequence 10878, A |
| 23 | 987 | 58.7 | 1149 | 54 | US-60-213-360-5290 | Sequence 107, App |
| 24 | 927.2 | 55.2 | 1131 | 54 | US-60-213-360-5290 | Sequence 5290, Ap |
| 25 | 856 | 50.9 | 1152 | 15 | US-09-171-562-3 | Sequence 3, Appli |
| 26 | 768.4 | 45.7 | 1606 | 29 | US-09-726-811-5246 | Sequence 5246, Ap |
| 27 | 756.8 | 45.0 | 969 | 58 | US-60-257-537-108 | Sequence 108, App |
| 28 | 661 | 39.3 | 790 | 30 | US-09-760-475-880 | Sequence 880, App |
| 29 | 638.4 | 38.0 | 1470 | 54 | US-60-213-360-5292 | Sequence 5292, Ap |
| 30 | 589.2 | 35.1 | 920 | 60 | US-60-278-258-10876 | Sequence 10876, A |
| 31 | 525.4 | 31.3 | 3360 | 60 | US-60-278-258-10880 | Sequence 10880, A |
| 32 | 521.8 | 31.0 | 2788 | 58 | US-60-257-537-114 | Sequence 114, App |
| 33 | 497.2 | 29.6 | 536 | 20 | US-09-535-896-29884 | Sequence 29884, A |
| 34 | 494.6 | 29.4 | 706 | 52 | US-60-192-739-1083 | Sequence 1083, Ap |
| 35 | 494.6 | 29.4 | 2207 | 30 | US-09-760-475-3851 | Sequence 3851, Ap |
| 36 | 494.6 | 29.4 | 2499 | 30 | US-09-760-475-3852 | Sequence 3852, Ap |
| 37 | 494.6 | 29.4 | 8427 | 30 | US-09-760-475-3849 | Sequence 3849, Ap |
| 38 | 494.6 | 29.4 | 8427 | 30 | US-09-760-475-3853 | Sequence 3853, Ap |
| 39 | 494.6 | 29.4 | 8427 | 30 | US-09-760-475-3857 | Sequence 3857, Ap |
| 40 | 490.8 | 29.2 | 510 | 20 | US-09-535-896-29803 | Sequence 29803, A |
| 41 | 488 | 29.0 | 3451 | 30 | US-09-770-174-4193 | Sequence 4193, Ap |

| | | | | | | |
|----|-------|------|------|----|----------------------|--------------------|
| 42 | 486.6 | 28.9 | 1145 | 50 | US-60-172-373-21412 | Sequence 21412, A |
| 43 | 459.8 | 27.4 | 611 | 58 | US-60-253-652-23792 | Sequence 23792, A |
| 44 | 432.6 | 25.7 | 439 | 25 | US-09-654-617-263882 | Sequence 263882, A |
| 45 | 432.6 | 25.7 | 439 | 27 | US-09-684-016-263882 | Sequence 263882, B |
| 46 | 432.6 | 25.7 | 439 | 27 | US-09-684-016-263882 | Sequence 263882, C |

ALIGNMENTS

```

RESULT      1
US-09-164-862B-4
; Sequence 4, Application US/09164862B
; GENERAL INFORMATION:
; APPLICANT: Price, Paul
; APPLICANT: Johansen, Julia
; TITLE OF INVENTION: YKL-40 AS A MARKER AND PROGNOSTIC INDICATOR FOR CANCERS
; FILE REFERENCE: 407T-8955-00US
; CURRENT APPLICATION NUMBER: US/09/164, 862B
; CURRENT FILING DATE: 1998-10-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 1681
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-164-862B-4

```

| | | | | |
|----------------------------|---------|--------------|--------|--------------|
| Query Match | 100.0%; | Score 1681; | DB 15; | Length 1681; |
| Best Local Similarity | 100.0%; | Pred. No. 0; | | |
| Matches 1681: Conservative | 0; | Mismatches | 0; | Indels 0; |
| Gaps | 0; | | | |

| | | | |
|----|-----|-----------------------------------------------------------------|-----|
| Qy | 1 | ctagtgagctggcacacgagcgctgggcaaggggaagaggccacacccctgcctgctctg | 60 |
| Db | 1 | ctagtgagctggcacacgagcgctgggcaaggggaagaggccacacccctgcctgctctg | 60 |
| Qy | 61 | ctgcagccagaaatgggtgtgaaggcgtctcaacacagggtttgtgtctcgtgtgtctcc | 120 |
| Db | 61 | ctgcagccagaaatgggtgtgaaggcgtctcaacacagggtttgtgtctcgtgtgtctcc | 120 |
| Qy | 121 | agtgctgctctgcatacaaaactgggtctgtatacaacagctgggtccacgtaccgggaag | 180 |
| Db | 121 | agtgctgctctgcatacaaaactgggtctgtatacaacagctgggtccacgtaccgggaag | 180 |
| Qy | 181 | gcgatggagcgtgctccacagatgcctctgacgcgtctctgtgtaccacacatcatctaca | 240 |
| Db | 181 | gcgatggagcgtgctccacagatgcctctgacgcgtctctgtgtaccacacatcatctaca | 240 |
| Qy | 241 | gctttgccaatataagcaacgatacatcgacacctgggagtggaatgatgtgacgtctct | 300 |
| Db | 241 | gctttgccaatataagcaacgatacatcgacacctgggagtggaatgatgtgacgtctct | 300 |
| Qy | 301 | acggcatgctcaacacactcaacacacgaaaccccaacctggaagactctctgtctgtctcg | 360 |
| Db | 301 | acggcatgctcaacacactcaacacacgaaaccccaacctggaagactctctgtctgtctcg | 360 |
| Qy | 361 | gaggatggaactttgggtctcaagaattttccaagatagacctccaaccccagagtcgcc | 420 |
| Db | 361 | gaggatggaactttgggtctcaagaattttccaagatagacctccaaccccagagtcgcc | 420 |
| Qy | 421 | ggactttcatcaagtcagtcgcgcattttctgcgacccatggctttgatggcgctgacc | 480 |
| Db | 421 | ggactttcatcaagtcagtcgcgcattttctgcgacccatggctttgatggcgctgacc | 480 |
| Qy | 481 | ttgccttggctctaaccttggcgggagagacaaacaccattttacaccccttaatacaggaaa | 540 |
| Db | 481 | ttgccttggctctaaccttggcgggagagacaaacaccattttacaccccttaatacaggaaa | 540 |
| Qy | 541 | tgaagccgaatttataaagggaagcccgccaggggaaaaagcagctcctctcagcgccag | 600 |
| Db | 541 | tgaagccgaatttataaagggaagcccgccaggggaaaaagcagctcctctcagcgccag | 600 |
| Qy | 601 | ctctctctctccacgaagctcaacattatcacacagctatgacatttgcacaaatatcccaac | 660 |

| | | | | |
|---|---|------|---------------------------------------------------------------|------|
| D | b | 601 | cactgtctgcgggaaagtcacattgcagcagctatgaccttgccagagtatcccac | 660 |
| Q | y | 661 | acctggattcattagcatcatcaactacgatatttcatggcgctggctgggaccacag | 720 |
| D | b | 661 | acctggattcattagcatcatgaactacgatatttcatggcgctggctgggaccacag | 720 |
| Q | y | 721 | gccatacagtcacctcaggcgaggctcaggaggatccaagtcctgcacagattcagcaaca | 780 |
| D | b | 721 | gccatacagtcacctcaggcgaggctcaggaggatccaagtcctgcacagattcagcaaca | 780 |
| Q | y | 781 | ctgactatgcttgtgggtacatgttaaggcttgggggctcctgcagtaagctggtgatgg | 840 |
| D | b | 781 | ctgactatgcttgtgggtacatgtttgaggctggggctcctgcagtaagctggtgatgg | 840 |
| Q | y | 841 | gcattccccacttcggagagagcttcaacttggctttcttgagactgggtttccagcgc | 900 |
| D | b | 841 | gcattccccacttcggagagagcttcaacttggctttcttgagactgggtttccagcgc | 900 |
| Q | y | 901 | caatctcaggaccgggaattccaggccgggttccaccaaggaggcaggacaccttgctact | 960 |
| D | b | 901 | caatctcaggaccgggaattccaggccgggttccaccaaggaggcaggacaccttgctact | 960 |
| Q | y | 961 | atgagatgttgacttcttcgcggagccacagfcccataagaacctcgcgcagcagctcc | 1020 |
| D | b | 961 | atgagatgttgacttcttcgcggagccacagfcccataagaacctcgcgcagcagctcc | 1020 |
| Q | y | 1021 | ccatatccaccaaggccaaccagttggttaggatacgcacacaggaaaagcgtcaaaagca | 1080 |
| D | b | 1021 | ccatatccaccaaggccaaccagttggttaggatacgcacacaggaaaagcgtcaaaagca | 1080 |
| Q | y | 1081 | aggtgcagtcacctgaagatagggcagctggcaggcgccatggtatggcgccctgacctgg | 1140 |
| D | b | 1081 | aggtgcagtcacctgaagatagggcagctggcaggcgccatggtatggcgccctgacctgg | 1140 |
| Q | y | 1141 | atgacttcaaggctccttctgcgccaggatctgcttccctctcaccaatgccatca | 1200 |
| D | b | 1141 | atgacttcaaggctccttctgcgccaggatctgcttccctctcaccaatgccatca | 1200 |
| Q | y | 1201 | aggatgcactgcgtgcacagctagccctctgttgcacacagcagcggggccaaggatgc | 1260 |
| D | b | 1201 | aggatgcactgcgtgcacagctagccctctgttgcacacagcagcggggccaaggatgc | 1260 |
| Q | y | 1261 | cccgctcccgcttgctggcgccggagccctgacactgccttgcctgcagtcgccaggtgag | 1320 |
| D | b | 1261 | cccgctcccgcttgctggcgccggagccctgacactgccttgcctgcagtcgccaggtgag | 1320 |
| Q | y | 1321 | ccctcagttccctcccttgggocctatgcagaggtcccaacacacagatttgagctcag | 1380 |
| D | b | 1321 | ccctcagttccctcccttgggocctatgcagaggtcccaacacacagatttgagctcag | 1380 |
| Q | y | 1381 | cccttggtyggcagagaggtacacacttgttgatgattaatgaaaagtgttacagatcccc | 1440 |
| D | b | 1381 | cccttggtyggcagagaggtacacacttgttgatgattaatgaaaagtgttacagatcccc | 1440 |
| Q | y | 1441 | aagcctggccaaggaaattctcaactccctgccccctagccctcttatcaaaggacac | 1500 |
| D | b | 1441 | aagcctggccaaggaaattctcaactccctgccccctagccctcttatcaaaggacac | 1500 |
| Q | y | 1501 | cattttggcagctctatcaccaggagccaaacatcctcaagaacacagtgacctact | 1560 |
| D | b | 1501 | cattttggcagctctatcaccaggagccaaacatcctcaagaacacagtgacctact | 1560 |
| Q | y | 1561 | aattatacccttcgaaagccagcttgaaccttcaacttaggaacgttaaactgtgtccct | 1620 |
| D | b | 1561 | aattatacccttcgaaagccagcttgaaccttcaacttaggaacgttaaactgtgtccct | 1620 |
| Q | y | 1621 | atcctaacttcccttccataattccacagctgctcaataaagtacaaagtttaacagtgt | 1680 |
| D | b | 1621 | atcctaacttcccttccataattccacagctgctcaataaagtacaaagtttaacagtgt | 1680 |
| Q | y | 1681 | g | 1681 |

Db 1316 ccggtcccccctctggtccagctggccggagcctgataccctgctgtagtcccaag 1375
Qy 1315 gctgagcctcagctccctcccttgggctcagaggtccacaacacagatttga 1374
Db 1376 gctgagcctcagctccctcccttgggctcagaggtccacaacacagatttga 1435
Qy 1375 gctcagcctctggtggcgagagag----- 1397
Db 1436 gctcagcctctggtggcgagaggttaggagtggtggtgggatagtgaggtcgcaaa 1495
Qy 1398 -----gtacacacttgtagtattgaaatgtttacagatcccc 1440
Db 1496 tgaagactcgggattaglacacacttgtagtattgaaatgtttacagatcccc 1555
Qy 1441 aagctggcaaggaaatttctcaactccctgcccctagcctccttatacaaaagacac 1500
Db 1556 aagctggcaaggaaatttctcaactccctgcccctagcctccttatacaaaagacac 1615
Qy 1501 catttggcaagctctatcaccaagagagccaaacatccttacaagacacagtgaccatact 1560
Db 1616 catttggcaagctctatcaccaagagagccaaacatccttacaagacacagtgaccatact 1675
Qy 1561 aattatacccccgtcaaacgagcagcttgaaacttcaacttaggaacgtaatcgtgtccct 1620
Db 1676 aattatacccccgtcaaacgagcagcttgaaacttcaacttaggaacgtaatcgtgtccct 1735
Qy 1621 atcctacttcccttccctaatccacagctgctcaataaaagtacaagagtttaacagtg 1680
Db 1736 atcctacttcccttccctaatccacagctgctcaataaaagtacaagagtttaacagtg 1795
Qy 1681 g 1681
Db 1796 g 1796

RESULT 5
US-60-257-537-115
; Sequence 115, Application US/60257537
; GENERAL INFORMATION:
; APPLICANT: Jones, Karen
; APPLICANT: Schafer, Alan
; TITLE OF INVENTION: NUCLEOTIDE POLYMORPHISMS ASSOCIATED WITH OSTEOARTHRITIS
; FILE REFERENCE: GX-0022 P
; CURRENT APPLICATION NUMBER: US/60/257,537
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 499
; SOFTWARE: PERL Program
; SEQ ID NO 115
; LENGTH: 1925
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; OTHER INFORMATION: CH13L1_mrna_build.1
US-60-257-537-115

Query Match 93.7%; Score 1575.4; DB 58; Length 1925;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1665; Conservative 0; Mismatches 16; Indels 60; Gaps 2;

Qy 1 ctaggtagctggcaccaggagccgtgggcaagggaaggagggccacaccctgcccctgctcg 60
Db 56 ctaggtagctggcaccaggagccgtgggcaagggaaggagggccacaccctgcccctgctcg 115
Qy 61 ctgagccgaagtgggtggaaggcgtctcaaacaggctttgttcctggtgctgctcc 120
Db 116 ctgagccgaagtgggtggaaggcgtctcaaacaggctttgttggctcctggctgctcc 175
Qy 121 agtgcgtcctgcatacaaacgtgctgtactacaccagctgggtcccagtagtccgggaag 180
Db 176 agtgcgtcctgcatacaaacgtgctgtactacaccagctgggtcccagtagtccgggaag 235

Qy 181 gcgatgggagctgcttccagatgccccttgaccgttctctgtgtaccacatcatctaca 240
Db 236 gcgatgggagctgcttccagatgccccttgaccgttctctgtgtaccacatcatctaca 295
Qy 241 gctttgcaatataagcaacgatcacatcgacacctgggagtggaatgtagtgcgtcct 300
Db 296 gctttgcaatataagcaacgatcacatcgacacctgggagtggaatgtagtgcgtcct 355
Qy 301 acgggatgctcaacacactcaacaacacagaaaccccaactgaagactctctgtgtctcg 360
Db 356 acgggatgctcaacacactcaagaacaggaaccccaactgaagactctctgtgtctcg 415
Qy 361 agggatgaaacttgggtctcaaaagattttccaaagatagcctcccaacccagagtcgcc 420
Db 416 agggatgaaacttgggtctcaaaagattttccaaagatagcctcccaacccagagtcgcc 475
Qy 421 ggaacttcaatcaagtcagtcacgcatttctcgcacccatgcttcttgatggcgctgacc 480
Db 476 ggaacttcaatcaagtcagtcacgcatttctcgcacccatgcttcttgatggcgctgacc 535
Qy 481 ttgcctggctacacctgacgagagacaaacacacattttaccacccctaatcaaggaaa 540
Db 536 ttgcctggctacacctgacgagagacaaacacacattttaccacccctaatcaaggaaa 595
Qy 541 tgaaggccgaatttataaaggagccagccaggggaaaaagcagctcctgctcagcgag 600
Db 596 tgaaggccgaatttataaaggagccagccaggggaaaaagcagctcctgctcagcgag 655
Qy 601 cactgtctcgggggaagggtcaacattgacagcagctatgacattgccaagatatccaac 660
Db 656 cactgtctcgggggaagggtcaacattgacagcagctatgacattgccaagatatccaac 715
Qy 661 acctgatttcattagcatcatgacctagattttcatgctgcctggctggggaccacag 720
Db 716 acctgatttcattagcatcatgacctagattttcatgctgcctggctggggaccacag 775
Qy 721 gccatcacagtcctccctcaggcgaggtcaggagagtcagagtcagctcctgacagatac 780
Db 776 gccatcacagtcctccctgttcagggtcagggtcagggtcagggtcagggtcagggtc 835
Qy 781 ctgaactatgctgtgggttacatgttgaggtcggggctcctcctccagtaagctggtgatg 840
Db 836 ctgaactatgctgtgggttacatgttgaggtcggggctcctcctccagtaagctggtgatg 895
Qy 841 qcattccacacttcgggagagcttccactctgcttctcttgagactggtgttccagcgc 900
Db 896 gcatcccccacttcgggagagcttccactctgcttctcttgagactggtgttggagccc 955
Qy 901 caatctcaggaccgggaattccaggccgggttccacaaaggaggaaggaccccttgcctact 960
Db 956 caatctcaggaccgggaattccaggccgggttccacaaaggaggaaggaccccttgcctact 1015
Qy 961 atgagatctgacttctccgcgagccacagtcctccatagaaacctcgccagcaggtcc 1020
Db 1016 atgagatctgacttctccgcgagccacagtcctccatagaaacctcgccagcaggtcc 1075
Qy 1021 cctatgccaccaaggccaaccagtggttaggatacgcacgcagcaggaagcgtcaaaagca 1080
Db 1076 cctatgccaccaaggccaaccagtggttaggatacgcacgcagcaggaagcgtcaaaagca 1135
Qy 1081 aggtgcagtaacctgaaggatagggcagctggcaggcgcctatgtagccctggacctgg 1140
Db 1136 aggtgcagtaacctgaaggatagggcagctggcaggcgcctatgtagccctggacctgg 1195
Qy 1141 atgaactccagggtccctctcgcggcaggatctgcgttccctctcccaatgcatca 1200
Db 1196 atgaactccagggtccctctcgcggcaggatctgcgttccctctcccaatgcatca 1255
Qy 1201 aggatgcactcgtcaacgtagcctctgttctgcacacagcagggggccaaagatgc 1260
Db 1256 aggatgcactcgtcgaacgtagcctctgttctgcacacagcagggggccaaagatgc 1315
Qy 1261 cccgtccccctgctg-----gctggccgggagcctgatcacctgcctgctgagtcctcag 1314


```
|||||
Db 1256 aaggtagcactgcgtgcaacgtagccctctgtcttgacacagcagcggggcccaaggatg 1315
Qy 1260 cccggtcccgctg-----gctggccggagcctgacacctccctgctgagtccca 1313
Db 1316 cccggtcccgctgctgcacgtgcccggagcctgacacctccctgctgagtccca 1375
Qy 1314 gctgagcctcagctccctccctgggcccctatgcagaggtcccaacacacacagatttg 1373
Db 1376 ggtgagcctcagctccctccctgggcccctatgcagaggtcccaacacacacagatttg 1435
Qy 1374 agtcagcctggtgggcagag----- 1397
Db 1436 agtcagcctggtgggcagaggttagggatgggctgtgggtagtgaggcatcgca 1495
Qy 1398 -----gtacacctattgattgattgaaatgaaatgtttacagatccc 1439
Db 1496 atgaagactcgggattgatacacattgtgatgattgaaatgtttacagatccc 1555
Qy 1440 caagcctggcaagggaattttcactcctgcctccctagcctccttatcaaaaggaca 1499
Db 1556 caagcctggcaagggaattttcactcctgcctccctgcctccttatcaaaaggaca 1615
Qy 1500 ccattttgcaagctctatcaccaaaggagccaaacatctctacaaagacacagtcacac 1559
Db 1616 ccattttgcaagctctatcaccaaaggagccaaacatctctacaaagacacagtcacac 1675
Qy 1560 taattataccctcgaagccagcttgaaaccttcaacttaggaacgtatgtctccc 1619
Db 1676 taattataccctcgaagccagcttgaaaccttcaacttaggaacgtatgtctccc 1735
Qy 1620 tatctacttcccttctaatctccacagctgtcctaataaagtacagagtttaacagt 1679
Db 1736 tatctacttcccttctaatctccacagctgtcctaataaagtacagagtttaacagt 1795
Qy 1680 tg 1681
Db 1796 tg 1797

RESULT 7
US-09-577-408-6988
; Sequence 6988, Application US/09577408
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Tillinghast, John
; APPLICANT: Sinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 792
; CURRENT APPLICATION NUMBER: US/09/577,408
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 8502
; SOFTWARE: pt_GCT_genes Version 1.0
; SEQ ID NO 6988
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (199)...(1198)
; OTHER INFORMATION: similar to q12286219 in the genepept database release 115,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-577-408-6988

Query Match 92.9%; Score 1561.2; DB 22; Length 1797;
Best Local Similarity 95.5%; Pred. No. 0;
Matches 1663; Conservative 0; Mismatches 18; Indels 61; Gaps 3;

Qy 1 ctaggtagtgccaccagagccgtgggcaagggaagggccacacccctgccctgctctg 60
```

```
|||||
Db 56 ctaggtagtgccaccagagccgtgggcaagggaagggccacacccctgccctgctctg 115
Qy 61 ctgagccagaatgggtgtgaagcgtctcaaacaggctttgtgtgctgtgtgtgtgtgtc 120
Db 116 ctgagccagaatgggtgtgaagcgtctcaaacaggctttgtgtgctgtgtgtgtgtc 175
Qy 121 agtgtgtctgtacatacaaaactgctctactacacacagctggcccagtlaccgggaag 180
Db 176 agtgtgtctgtacatacaaaactgctctactacacacagctggcccagtlaccgggaag 235
Qy 181 gcgagggagctgtccocagatgcccttgaccgttctctgtgtacccacatcatctaca 240
Db 236 gcgagggagctgtccocagatgcccttgaccgttctctgtgtacccacatcatctaca 295
Qy 241 gcttgccaataataagcaacatcatcatcgacacctgggagtggaatgatgacgctct 300
Db 296 gcttgccaataataagcaacatcatcatcgacacctgggagtggaatgatgacgctct 355
Qy 301 acggcatgctcaacacactcaaacacacgaaaccccaacctgaagactctctgtgtgtcg 360
Db 356 acggcatgctcaacacactcaaacacacgaaaccccaacctgaagactctctgtgtgtcg 415
Qy 361 gaggatggaaactttgggtctcaaaagattttccaagatagcctccaaccccagatcgcc 420
Db 416 gaggatggaaactttgggtctcaaaagattttccaagatagcctccaaccccagatcgcc 475
Qy 421 ggactttcatcaagtcagtcacgcctttctgcgaccccatggcttggatggcgtagacc 480
Db 476 ggactttcatcaagtcagtcacgcctttctgcgaccccatggcttggatggcgtagacc 535
Qy 481 ttgcctggctctacctggacgagagacaaacacacattttaccacccctaaatacaaggaaa 540
Db 536 ttgcctggctctacctggacgagagacaaacacacattttaccacccctaaatacaaggaaa 595
Qy 541 tgaagcccgaaatttataaaggaaagccagccaggggaaaaagcagctcctgcagcgcag 600
Db 596 tgaagcccgaaatttataaaggaaagccagccaggggaaaaagcagctcctgcagcgcag 655
Qy 601 cactgtctcggggaaaggtcacccattgacagcagctatgacattgccaagatcccaac 660
Db 656 cactgtctcggggaaaggtcacccattgacagcagctatgacattgccaagatcccaac 715
Qy 661 acctgatttcatttagcatcatgacctacgattttcattggcctcctggcgtgggaccacag 720
Db 716 acctgatttcatttagcatcatgacctacgattttcattggcctcctggcgtgggaccacag 775
Qy 721 gccatcacagtcctccctcagcgcgaggtcaggagatgcaagtctctgacagattcagaaca 780
Db 776 gccatcacagtcctccctcagcgcgaggtcaggagatgcaagtctctgacagattcagaaca 835
Qy 781 ctgactatgctgtgggtacatgttgagggtcggggctcctcctgcccagtaagctggtgatgg 840
Db 836 ctgactatgctgtgggtacatgttgagggtcggggctcctcctgcccagtaagctggtgatgg 895
Qy 841 gcattcccaacctcgggagagcttcaactcctgttctctgtgagactgggttccagcgc 900
Db 896 gcattcccaacctcgggagagcttcaactcctgttctctgtgagactgggttccagcgc 955
Qy 901 caatctcaggaccgggaattccaggccggttccacaaaggagggagggacccttgcctact 960
Db 956 caatctcaggaccgggaattccaggccggttccacaaaggagggagggacccttgcctact 1015
Qy 961 atgagatgtgacttctcctccgagagccacagtcctatagacccctcggccagcaggtccc 1020
Db 1016 atgagatgtgacttctcctccgagagccacagtcctatagacccctcggccagcaggtccc 1075
Qy 1021 cctatgccaccaagggcaaccagtggtgtagatagcagcagcaggaagagcgtcaaaagca 1080
Db 1076 cctatgccaccaagggcaaccagtggtgtagatagcagcagcaggaagagcgtcaaaagca 1135
Qy 1081 aggtgtca-gtacctgaagatagcagctggcagcgcctatggatggcctgagacctg 1139
|||||
```


[illegible]

|||||
Db 1239 atgactccaggctcctctctcgccgagatctcgcttccctctcaccatgccaatca 1298
QY 1201 aggatcactcgtcgaaagtagcctctgttctgcacacagacacgagggccaaagatgc 1260
Db 1299 aggatcactcgtcgaaagtagcctctgttctgcacacagacacgagggccaaagatgc 1358
QY 1261 cccgtcccgctctg-----gctggccgagagcctgatcacctgcctcgtgagtcaccag 1314
Db 1359 cccgtcccgctctgctccagctggtcgggagcctgatcacctgcctcgtgagtcaccag 1418
QY 1315 gctgagcctcagtcctccctcccttgggacctatgcagaggtccacacacacagatttga 1374
Db 1419 gctgagcctcagtcctccctcccttgggacctatgcagaggtccacacacacagatttga 1478
QY 1375 gctcagcctcgtggcgacagaggtacacacttg----- 1408
Db 1479 gctcagcctcgtggcgacagaggtaggatgggctgtgggagtagtgaggatcgcaa 1538
QY 1409 -----ttgatgattaatgaaatgtttacagatccccaag 1443
Db 1539 tgaagactcgggattagtagcacactgttga ttaatggaaatgtttacaga tccccaag 1598
QY 1444 cctggcaagggaattttctcaactcctcgtccctcctagccctcttatcaaaagacacccat 1503
Db 1599 cctggcaagggaattttctcaactcctcgtccccaagcctccttatcaaaagacacccat 1658
QY 1504 ttggcaagctctatcacaaaggagcaaacatctctacaaagacacagtgaccatactaat 1563
Db 1659 ttggcaagctctatcacaaaggagcaaacatctctacaaagacacagtgaccatactaat 1718
QY 1564 tatcccccctgcaag-ccaagcttgaacctcacttaggaacgtaatcgtgtccctat 1622
Db 1719 tataccccctgcaaaagcctgtgaaccttgaaccttcaacttaggaacgtaatcgtgtccctat 1778
QY 1623 cctacttcccccttcaattccacagctgctcaataaaagtacaagatttaacagtggtg 1691
Db 1779 cctacttcccccttcaattccacagctgctcaataaaagtacaagatttaacagtggtg 1837

RESULT 10
US-60-164-285-3401
; Sequence 3401, Application US/60164285
; GENERAL INFORMATION:
; APPLICANT: Ma, Xiao-Jun
; TITLE OF INVENTION: Tumor Associated Molecules (TAMs): Targets for diagnosis, treatment
; FILE REFERENCE: 3214
; CURRENT APPLICATION NUMBER: US/60/164,285
; CURRENT FILING DATE: 1999-11-05
; NUMBER OF SEQ ID NOS: 8259
; SEQ ID NO 3401
; LENGTH: 1739
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: unsure at all n locations
US-60-164-285-3401

Query Match 91.7%; Score 1542; DB 49; Length 1739;
Best Local Similarity 94.9%; Pred. No. 0;
Matches 1648; Conservative 0; Mismatches 31; Indels 58; Gaps 3;

QY 1 ctaggtagctggccacagagccgtggcaagggaagagccacacccctgctctg 60
Db 3 ctaggtagctggccacagagccgtggcaagggaagagccacacccctgctctg 62
QY 61 ctgcagccagaatgggtgtgaaggcgtctcaaacagggcttgggtcctggtgctctcc 120
Db 63 ctgcagccagaatgggtgtgaaggcgtctcaaacagggcttgggtcctggtgctctcc 122
QY 121 agtgcgtcgtcgcatacaactggtctgtactacaccagctggtcccaagcgggaag 180
|||||

Db 123 agtgcgtcgtcgcatacaaaaactggtctgtctactacaccagctggtcccaagtcacgggaag 182
QY 181 gcatggggagctgcttcccagatgccttgcagcgttctctgtgtaccacatcatctaca 240
Db 183 gcatggggagctgcttcccagatgccttgcagcgttctctgtgtaccacatcatctaca 242
QY 241 gctttgccataataagcaacgatacacatgcacacccctgggagtggaatgatgtgacgctct 300
Db 243 gctttgccataataagcaacgatacacatgcacacccctgggagtggaatgatgtgacgctct 302
QY 301 acggcatgctcaaacacactcaacaacacaccccccacccctgaagactctcttctgtctg 360
Db 303 acggcatgctcaaacacactcaagaacacaggaaccccaacccctgaagactctcttctgtctg 362
QY 361 gaagatggaaacttgggtctcaaaagattttccaagatagctcctcaaacacacagagtcgc 420
Db 363 gaagatggaaacttgggtctcaaaagattttccaagatagctcctcaaacacacagagtcgc 422
QY 421 ggactttcatcaagtcagtcaccgccatttctgcgaccccatggtcttgcagtcgtgac 480
Db 423 ggactttcatcaagtcagtcaccgccatttctgcgaccccatggtcttgcagtcgtgac 482
QY 481 ttgcctggctctaccctggacggagagacaaacacacattttaccacccctaatcaaggaaa 540
Db 483 ttgcctggctctaccctggacggagagacaaacacagcattttaccacccctaatcaaggaaa 542
QY 541 tgaagccgaatttataaaggaaagccagccaggggaaaaagcagctcctctcagcgacg 600
Db 543 tgaagccgaatttataaaggaaagccagccaggggaaaaagcagctcctctcagcgac 602
QY 601 cactgtctcgggggaaggtcacattgacagcagctatgacattgccaagatatacccaac 660
Db 603 acntgtctcgggggaaggtcacattgacagcagctatgacattgccaagatatacccaac 662
QY 661 acctgatttcaatagcatcatcactacatgatttctatggcctggtggaccacag 720
Db 663 acctgatttcaatagcatcatcactacatgatttctatggcctggtggaccacag 722
QY 721 gccatcacagtcctccagcaggtcagaggtcagaggtcgaagctcgcagattcagacgaaca 780
Db 723 gccatcacagtcctccagcaggtcagaggtcagaggtcgaagctcgcagattcagacgaaca 782
QY 781 ctgactatgctgtgggtacatgttggagctgggggctcctcgcagtaagctggtgatgg 840
Db 783 ctgactatgctgtgggtacatgttggagctgggggctcctcgcagtaagctggtgatgg 842
QY 841 gcataccacacttcgggagagctcactcgtcttcttcttgagactggtgttccacgcg 900
Db 843 gcataccacacttcgggagagctcactcgtcttcttcttgagactggtgttggagccc 902
QY 901 caatctcaggacccgggaattccagccggttccaccaaggaggcagggacccttgcctact 960
Db 903 caatctcaggacccgggaattccagccggttccaccaaggaggcagggacccttgcctact 962
QY 961 atgagatctgtgacttctccgcggagccacagtcacatagaacctcgcgcagcaggtcc 1020
Db 963 atgagatctgtgacttctccgcggagccacagtcacatagaacctcgcgcagcaggtcc 1022
QY 1021 cctatgccaccaaggggcaaccagtggttaggtacacagcaccagggaagcgtcaaaagca 1080
Db 1023 cctatgccaccaaggggcaaccagtggttaggtacacagcaccagggaagcgtcaaaagca 1082
QY 1081 aggtgcagttacctgaaggatagtcagctggcagcgcacatgggtatggccctggacctgg 1140
Db 1083 aggtgcagttacctgaaggatagtcagctggcggcgccatgggtatggccctggacctgg 1142
QY 1141 atgacttccagggctccttctcggccaggtatctgcgttccctctcacaatgcacatca 1200
Db 1143 atgacttccagggctccttctcggccaggtatctgcgttccctctcacaatgcacatca 1202
QY 1201 aggatgcactcgtgcacacgtagccctctgttctgcacacagcagcggggcccaaggtgc 1260
Db 1203 aggatgcactcgtgcacacgtagccctctgttctgcacacagcagcggggcccaaggtgc 1262

Qy 1395 gaggtacacactgttgatgattgaaatgtttacagatccccaaagcctggcaagg 1454
|||||
Db 1261 gaggtacacactgttgatgattgaaatgtttacagatccccaaagcctggcaagg 1320
|||||
Qy 1455 aatttcttaactcctgcccctagcctctctatcaaaagacaccattttggcaagct 1514
|||||
Db 1321 aatttcttaactcctgcccctagcctctctatcaaaagacaccattttggcaagct 1380
|||||
Qy 1515 ctatcaccaaggagcaaacatctctacaagacacagtgaccatactaattataccccctg 1574
|||||
Db 1381 ctatcaccaaggagcaaacatctctacaagacacagtgaccatactaattataccccctg 1440
|||||
Qy 1575 caaaagcagctgaaaccttcaacttagaagcgtaatcgtgtccccctacttccccct 1634
|||||
Db 1441 caaaagcagctgaaaccttcaacttagaagcgtaatcgtgtccccctacttccccct 1500
|||||
Qy 1635 tctaattccacagctgtctcaataaagatacaagagtttaacagtgtg 1681
|||||
Db 1501 tctaattccacagctgtctcaataaagatacaagagtttaacagtgtg 1547
|||||
RESULT 12
US-09-577-408-6987
; Sequence 6987, Application US/09577408
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Tillinghast, John
; APPLICANT: Sinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 792
; CURRENT APPLICATION NUMBER: US/09/577,408
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 8502
; SOFTWARE: BLAST_Genes Version 1.0
; SEQ ID NO 6987
; LENGTH: 1547
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (10)...(1007)
; OTHER INFORMATION: similar to g1286219 in the genepept database release 115,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-577-408-6987

Query Match 91.6%; Score 1540.6; DB 22; Length 1547;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1543; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 135 tacaactgtgtctactacaccagctgttccagtagccaggaagcgatggagctgc 194
|||||
Db 1 tacaactgtgtctactacaccagctgttccagtagccaggaagcgatggagctgc 60
|||||
Qy 195 ttccagatgcccctgaccgcttctgtgtaccacatcatctacagctttgccaatata 254
|||||
Db 61 ttccagatgcccctgaccgcttctgtgtaccacatcatctacagctttgccaatata 120
|||||
Qy 255 agcaacgatcacatgacacacctggagtggaatgatgtgacgctctacggcatgtccaac 314
|||||
Db 121 agcaacgatcacatgacacacctggagtggaatgatgtgacgctctacggcatgtccaac 180
|||||
Qy 315 acactcaacaacagcaacccacctgaagactcttctgtcggagatggaacttt 374
|||||
Db 181 acactcaacaacagcaacccacctgaagactcttctgtcggagatggaacttt 240
|||||
Qy 375 ggggtctcaagatcttccaaagtagctctcaacacacagagtcgcccggactttcatcaag 434
|||||
Db 241 ggggtctcaagatcttccaaagtagctctcaacacacagagtcgcccggactttcatcaag 300
|||||

Qy 435 tcagtcacgcgcatttctgcgaccatggctttgatggcggtgaccttgctgctctac 494
|||||
Db 301 tcagtcacgcgcatttctgcgaccatggctttgatggcggtgaccttgctgctctac 360
|||||
Qy 495 cctggacggagagacaaacacacctttaccacccttaatacaaggaaatgaagccgaattt 554
|||||
Db 361 cctggacggagagacaaacacacctttaccacccttaatacaaggaaatgaagccgaattt 420
|||||
Qy 555 ataaaggaagccagccaggggaaaaagcagctccctgcacgagcagcactgctcgggg 614
|||||
Db 421 ataaaggaagccagccaggggaaaaagcagctccctgcacgagcagcactgctcgggg 480
|||||
Qy 615 aaggtcacaccattgacagcagctatgacattgccaagatatcccaacacctgatttcatt 674
|||||
Db 481 aaggtcacaccattgacagcagctatgacattgccaagatatcccaacacctgatttcatt 540
|||||
Qy 675 agcatatgacctacgattttcatggcgctggcgctgggaccacagggccatcacagtcctc 734
|||||
Db 541 agcatatgacctacgattttcatggcgctggcgctgggaccacagggccatcacagtcctc 600
|||||
Qy 735 ctcaagcgaggtcaggagagatgcaagtcctgcagatttcagcaacacactgactgtgtg 794
|||||
Db 601 ctgttccgaggtcaggagagatgcaagtcctgcagatttcagcaacacactgactgtgtg 660
|||||
Qy 795 ggttacatgttgagctggcgctcctgcagtaagctggtgatggcattccccaccttc 854
|||||
Db 661 ggttacatgttgagctggcgctcctgcagtaagctggtgatggcattccccaccttc 720
|||||
Qy 855 gggagagcttcaacttggcttcttgagactggtgttccagcgccaatctcagagaccg 914
|||||
Db 721 gggagagcttcaacttggcttcttgagactggtgttccagcgccaatctcagagaccg 780
|||||
Qy 915 ggaattccagcggttcaccagagggcaggagaccttgcctactatgatagtctgtgac 974
|||||
Db 781 ggaattccagcggttcaccagagggcaggagaccttgcctactatgatagtctgtgac 840
|||||
Qy 975 ttccctcggcgagccacagctccatagaaacctcggccagcaggttcccttatgccacaag 1034
|||||
Db 841 ttccctcggcgagccacagctccatagaaacctcggccagcaggttcccttatgccacaag 900
|||||
Qy 1035 ggcaaccagtggttagtatacagacaccaggaaagcgtcaaaagcaagtgtagtacctg 1094
|||||
Db 901 ggcaaccagtggttagtatacagacaccaggaaagcgtcaaaagcaagtgtagtacctg 960
|||||
Qy 1095 aaggtatgacgctggcagggcccatggtatggcgctggacctgagatgacttcccgggc 1154
|||||
Db 961 aaggtatgacgctggcagggcgccatggtatggcgctggacctgagatgacttcccgggc 1020
|||||
Qy 1155 tcttctcggcgagagatctgcgcttccctctcaccaatgccatcaagatgcaactcgct 1214
|||||
Db 1021 tcttctcggcgagagatctgcgcttccctctcaccaatgccatcaagatgcaactcgct 1080
|||||
Qy 1215 gcaacgtagccctctgttgcacacagcagcggggcccagagatgcccgctccccctgtg 1274
|||||
Db 1081 gcaacgtagccctctgttgcacacagcagcggggcccagagatgcccgctccccctgtg 1140
|||||
Qy 1275 gctggcgggagcctgatacactgcctctgtagtccccagggctgagcctcaagtctccctc 1334
|||||
Db 1141 gctggcgggagcctgatacactgcctctgtagtccccagggctgagcctcaagtctccctc 1200
|||||
Qy 1335 ccttggggcctatgagaggtccacaacacagatattgagctcagccctgggtgggcaga 1394
|||||
Db 1201 ccttggggcctatgagaggtccacaacacagatattgagctcagccctgggtgggcaga 1260
|||||
Qy 1395 gaggtacacactgttgatgattaaatgtttacagatccccaaagcctggcgaagg 1454
|||||
Db 1261 gaggtacacactgttgatgattaaatgtttacagatccccaaagcctggcgaagg 1320
|||||
Qy 1455 aatttctcaactcccccttagccctctcttatacaaggacacattttggcgaagct 1514
|||||
Db 1321 aatttctcaactcccccttagccctctcttatacaaggacacattttggcgaagct 1380
|||||

Qy 1515 ctatcaccaagagccaaacatctctacaagacacacagtgtaacatactattatccccctg 1574
Db 1381 ctatcaccaagagccaaacatctctacaagacacacagtgtaacatactattatccccctg 1440
Qy 1575 caaagccagcttgaaacctcaacttagaagcgaatgctgccccctacttccccct 1634
Db 1441 caaagccagcttgaaacctcaacttagaagcgaatgctgccccctacttccccct 1500
Qy 1635 tcttaattccacagctgctcaataaagtaacagaggttaacagtg 1681
Db 1501 tcttaattccacagctgctcaataaagtaacagaggttaacagtg 1547

RESULT 13
US-09-359-922-6683/c
; Sequence 6683. Application US/09359922
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONFIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-752CON1
; CURRENT APPLICATION NUMBER: US/09/359,922
; CURRENT FILING DATE: 1999-07-22
; EARLIER APPLICATION NUMBER: US 09/205,155
; EARLIER FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 13203
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6683
; LENGTH: 1786
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-359-922-6683

Query Match 91.4%; Score 1536.8; DB 17; Length 1786;
Best Local Similarity 95.1%; Pred. No. 0;
Matches 1652; Conservative 0; Mismatches 27; Indels 59; Gaps 4;

Qy 1 ctaggtagctggcaccagggcggcgtgggcaaggagagccacacccctgctctg 60
Db 1754 CTAGGTAGCTGGCACCAGGAGCGGTGGGCAAGGAGAGAGGCCACACCCCTGCTCTG 1695

Qy 61 ctgcagccagaatgggtgtgaaggcgtctcaaacagccttgctgctgctgctcc 120
Db 1694 CTGCAGCCAGAAATGGGTGTGAAGCGCTCTCAACAGCGCTTGTGGTCTGCTGCTCC 1635

Qy 121 agtctgctctgatacaaaactgtctgtactacacacagctggtccacgtaccggaa 180
Db 1634 AGTCTGCTCTGCATACAAACTGTCTGTCTACTACACACAGCTGCTCCAGTACCGGGA 1575

Qy 181 gcgatggagctgtctccagatgcccttgaccgtcttctgtgtaccacacatctaca 240
Db 1574 GCGATGGAGCTGCTCCAGATGCCCTTGACCGCTTCTCTGTACCCATCATCTACA 1515

Qy 241 gctttgccaataaagcaagatcacatcgacacccctggagtggaatgatgtgctct 300
Db 1514 GCTTTGCCAATAAAGCAACGATCACATCGACACCTGGGAGTGGAAATGATGTACGCTCT 1455

Qy 301 acgcatgctcaacacactcaaacacgaaccccaacctgaagactcttctgtgctg 360
Db 1454 ACGCATGCTCAACACACTCAAGAACAGGAACCCCAACCTGAAGACTCTCTGTCTGTG 1395

Qy 361 gagatgaaactttgggtctcaagatctttccaagatagctcccaacacccagagtcg 420
Db 1394 GAGATGGAACCTTTGGGTCTCAAGATTTTCCAGATAGCTCCCAACACCCAGAGTCGCC 1335

Qy 421 ggaactttcatcaagtacgtacgcgaattcttgcgacccatgggtttgatgggtgacc 480
Db 1334 GGACTTTTCATCAAGTCAGTACCGCATTTCTCGCACCCATCGGCTTTGATGGCTGGACC 1275

Qy 481 ttgcctgctctacccctggagagagacaacacacattttaccacccctaatcaaggaa 540
Db 1274 TTGCCTGCTCTACCCCTGGAGGAGAGACAACAGCATTTTACCACCCCTAATCAAGGAA 1215

Qy 541 tgaagccgaattataaagaa-gccagccagggaaaaagcagctcctgctcagcgca 599
Db 1214 TGAAGCCGAATTTATAAAGGAAGGCCAGCCAGCAGGAAAAAGCAGCTCTGCTCAGCGCA 1155

Qy 600 gcactgtctgggggaaagtcaccattgacagcagctatgacattgccaagatatcccaa 659
Db 1154 GCACGTCTGCGGGGAAGGTACCATTTGACAGCAGTATGACATTTGCCAAGATATCCCAA 1095

Qy 660 cacttgattcatttagcatcatgacctacgatttcatggcgccctggcgtgggaccaca 719
Db 1094 CACCTGGATTTTCATTAGCATCATGACCTACGATTTTCATGGAGCCTGGCGTGGGACCACA 1035

Qy 720 ggcacacagtcacctcagcgaggtcagggatgcaagtcctctgacagattcagcaac 779
Db 1034 GGCCATCACAGTCCCTCTGTTCCGAGGTGAGGAGTGAAGTCTCTGACAGATTCAGCAAC 975

Qy 780 actgactatgctgtgggttacatgttgaggtcggggcctcctgcccagtaagctggatg 839
Db 974 ACTGACTATGCTGTGGGTACATGTTGAGGCTGGGGCCTCTGCCAGTAAGCTGGTGATG 915

Qy 840 ggcacccacacccctggggagagcttcaactctgcttcttctgagactggtgtccagcg 899
Db 914 GGCATCCCACTTCGGGAGGAGCTTCACTCTGGCTTCTTCTGAGACTGGTGTGGAGCC 855

Qy 900 ccaatctcagaccgggaattccagcccggttccaccaaggagcagggaccccttgctac 959
Db 854 CCAATCTCAGGACCGGGAATTCAGGCCGGTTCCACCAAGGAGGACGGACCCCTTGCTAC 795

Qy 960 tatgagactgtgacttctcccgaggccacagtcocatagaacctctggcagcgaggtc 1019
Db 794 TATGAGACTGTGACTTCTCTCCCGGAGCCACAGTGCATAGAAATCTCTCGGCCAGCGTC 735

Qy 1020 ccctatgccaaccaaggaaccaagtggtgtagatagcagcagcagcaagaaagcgtcaaaagc 1079
Db 734 CCCTATGCCACCAAGGCCAACAGTAGGTAGGTAGGATACACACAGGAGGAAACGCTCAAAAGC 675

Qy 1080 aaggtcagctacatgaagtagtagcagctggcgagcgcccatgtagggccctggagactg 1139
Db 674 AAGTGCAGTACCTGAGGACAGCAGCTGGCGGCCCATGTGTATGGGCCCTTGGACCTG 615

Qy 1140 gatgaactccagggtcctcttcggcgccaggtatgagcttccctctcacaatgcatc 1199
Db 614 GATGACTTCCAGGCTCCTTCTGCGGCCAGGATCTGCGCTTCCCTCTCACAAATGCCATC 555

Qy 1200 aagatgacactgctgcaacgtagcctctgtctgcacagcacggtgggccaagatg 1259
Db 554 AAGGATGACTCGCTGCAACGTAGCCCTCTGTTCTGCACACAGCACCGGGGCCAAGGATG 495

Qy 1260 cccgctccccgtctg-----gctggcgggagcctgatcacctggcctgctgagtcoca 1313
Db 494 CCCCCTCCCCCTCTGGCTCCAGCTGGCGGGAGCCTGATCACCTGCCCTGCTGAGTCCCA 435

Qy 1314 ggctgagcctgaagtctccctccttggggcctatgcaagaggtccacaacacagatttg 1373
Db 434 GGCTGAGCCTCAGTCTCCCTCCCTTGGGGCCTATGCAAGAGGTCCACACACAGATTTG 375

Qy 1374 agctcagccctggtggcgagaggtacacacttg----- 1408
Db 374 AGCTCAGCCCTGGTGGGCAGAGGTAGGGATGGGGTGTGGGATAGTAGGCGATCCCA 315

Qy 1409 -----ttgatgattaaaggaaatgtttacagatcccaa 1442
Db 314 ATGTAAGACTCGGGAATTAGTACACACTTGTGTGATTAATGGAATGTTTACAGATCCCAA 255

Qy 1443 gccttggcaagggaatttcttcaactcctcctggcccttagccctccttatacaaggacaca 1502
Db 254 GCTTGGCAAGGGAATTTCTCAACTCCTCTGCCCGCCAGCCCTCTTATCAAAAGGACACA 195

Qy 1503 tttyggcaagctctatcaccaagggacaaactcactcaagacagactgaccataactaa 1562
Db 194 TTTTGGCAAGCTCTATCACCAGGAGGCCAAACATCTCTACAAAGACAGTGTGACCATACTAA 135

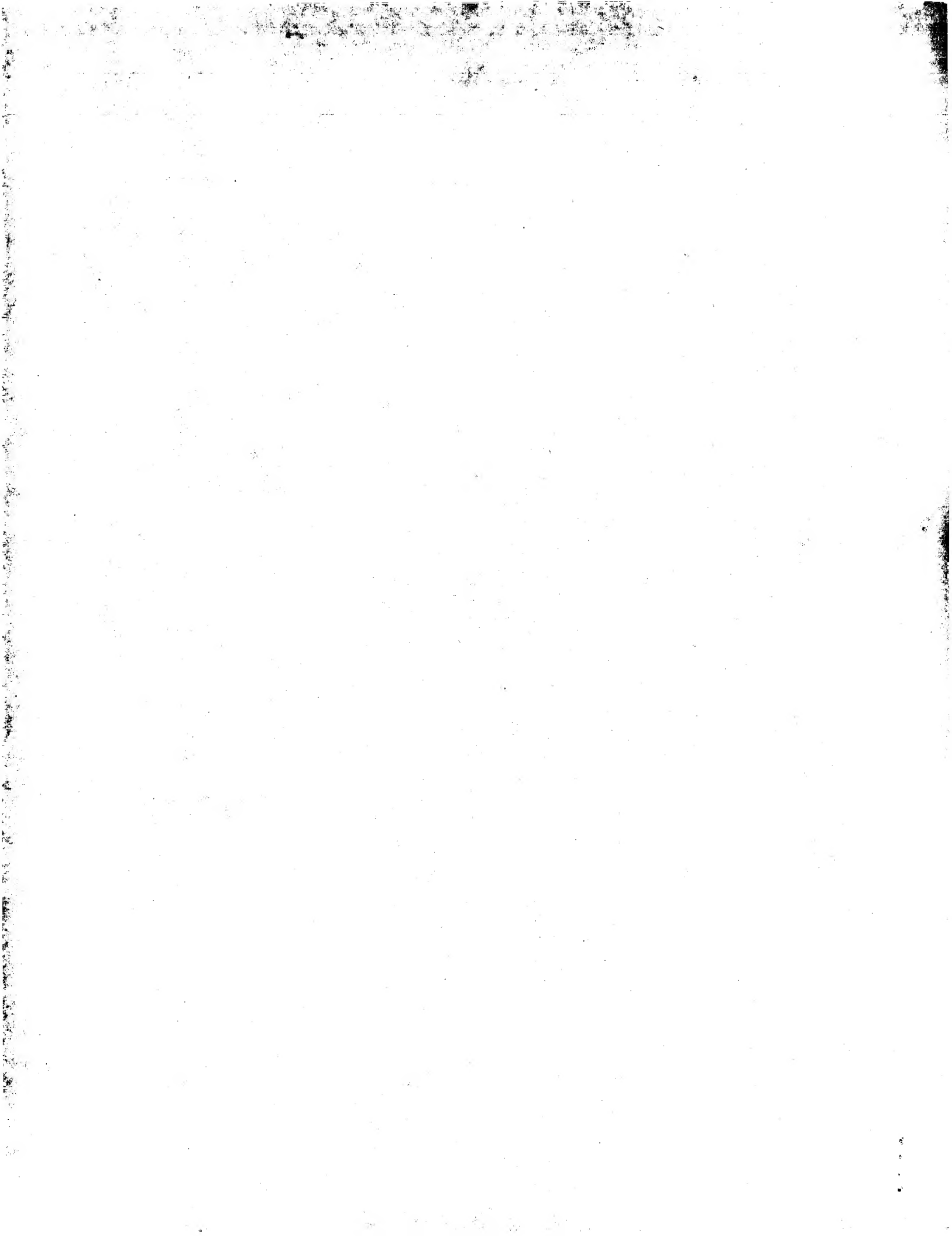
Db 134 TTATACCCCTGCAAAAGCCAGCTTGAAACCTTCACCTTAGGAACGTAATCGTGTCCCCCTA 75
Qy 1622 tctacttcccttcttaattccacagctgctcaataaagtacaaagatttaacagtg 1679
Db 74 TCCTACTTCCCTTCTTAATTCACAGCTGCTCAATAAAGTACAAGAGCTTAACAGTG 17
RESULT 15
US-60-278-258-10870
; Sequence 10870, Application US/60278258
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0010-1 P
; CURRENT APPLICATION NUMBER: US/60/278,258
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 17730
; SOFTWARE: PERL Program
; SEQ ID NO 10870
; LENGTH: 1824
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 237999.12
US-60-278-258-10870

Query Match 90.5%; Score 1521.2; DB 60; Length 1824;
Best Local Similarity 94.2%; Pred. No. 0;
Matches 1651; Conservative 0; Mismatches 28; Indels 73; Gaps 4;
Qy 1 ctaggtagctggcaccaggagccctgggcaaggaaagagccacacccctgccctgctctg 60
Db 70 ctaggtagctggcaccaggagccctgggcaaggaaagagccacacccctgccctgctctg 129
Qy 61 ctgagccagaatgggtgtaaggcgtctcaaacagagcctttgtggtcctggtgctgctc 120
Db 130 ctgagccagaatgggtgtaaggcgtctcaaacagagcctttgtggtcctggtgctgctc 189
Qy 121 agtgcctctgcatacaaaactgctgctactacacacagctggtccagtaaccggaag 180
Db 190 agtgcctctgcatacaaaactgctgctactacacacagctggtccagtaaccggaag 249
Qy 181 gcgatggagctgcttccagatgcccctggaccgcttctctgtaccacatcatctaca 240
Db 250 gcgatggagctgcttccagatgcccctggaccgcttctctgtaccacatcatctaca 309
Qy 241 gctttgccaataaagcaacgatacacatcgacacccctggagtggaatgatgtgacgtct 300
Db 310 gctttgccaataaagcaacgatacacatcgacacccctggagtggaatgatgtgacgtct 369
Qy 301 acggcatctcaacacatcaacacagaaaccccaacccctgaagactctctgtctgctg 360
Db 370 acggcatctcaacacatcaacagaaagaaaccccaacccctgaagactctctgtctgctg 429
Qy 361 gaggatggaactttgggtctca- -----aagattttccaagatagcctcca 405
Db 430 gaggatggaactttgggtctcaaaaggctgattttccaagattttccaagatagcctcca 489
Qy 406 acaccagagtcgcccagatttcaatcagtaagtaaccgcccattttctgcgacccatgact 465
Db 490 acaccagagtcgcccagatttcaatcagtaagtaaccgcccattttctgcgacccatgact 549
Qy 466 ttgatggcgtagaccttgcctgctctaccctggagcagagacaaacacattttacca 525
Db 550 ttgatggcgtagaccttgcctgctctaccctggagcagagacaaacacattttacca 609

Qy 526 ccctaatacaagaaatgaagcccgaaattataaaggaagccagccagggaaaaagcagc 585
Db 610 ccctaatacaagaaatgaagcccgaaattataaaggaagccagccagggaaaaagcagc 669
Qy 586 tctgtcagcgcagcactgtctgcggggaaggtcaccaattgacagcagctatgacattg 645
Db 670 tctgtcagcgcagcactgtctgcggggaaggtcaccattgacagcagctatgacattg 729
Qy 646 ccaagatatcccaacacactggatttcatttagcatcatcaccatcatttccatgctgct 705
Db 730 ccaagatatcccaacacactggatttcatttagcatcatcaccatcatttccatgctgct 789
Qy 706 ggcgtgggaccacagcagcaccatcagctccctcagcgaggtcagaggaagtcgaagtcctg 765
Db 790 ggcgtgggaccacagcagcaccatcagctccctcagcgaggtcagaggaagtcgaagtcctg 849
Qy 766 acagattcagcaacactgactgtctgtgggtacatgtgtgaggtcggggctcctgcaca 825
Db 850 acagattcagcaacactgactgtctgtgggtacatgtgtgaggtcggggctcctgcaca 909
Qy 826 gtaagctggtgatgggcacccacactcgggagagctcactctggtctctcttgaga 885
Db 910 gtaagctggtgatgggcacccacactcgggagagctcactctggtctctcttgaga 969
Qy 886 ctggtgtccagcgccaactcagggaccgggaattccagcggttcacaaaggagcgag 945
Db 970 ctggtgtggagcccaactcagggaccgggaattccagcggttcacaaaggagcgag 1029
Qy 946 ggaccttgctactatgatgactgtgacttctcgcggagccacagtcacatagaacc 1005
Db 1030 ggaccttgctactatgatgactgtgacttctcgcggagccacagtcacatagaacc 1089
Qy 1006 tcggcagcaggtccctctatgccacaaaggccacagtggtgtaggatacgacagcag 1065
Db 1090 tcggcagcaggtccctctatgccacaaaggccacagtggtgtaggatacgacagcag 1149
Qy 1066 aaagcgtcaaaagcaaggtcagctacctgaagtagagcagctggcaggcgccatggtat 1125
Db 1150 aaagcgtcaaaagcaaggtcagctacctgaagtagagcagctggcaggcgccatggtat 1209
Qy 1126 gggccctggacctggatgacttccagggctcctctcgcggcaggatctgcgttccctc 1185
Db 1210 gggccctggacctggatgacttccagggctcctctcgcggcaggatctgcgttccctc 1269
Qy 1186 tcaccaatgcatcaaaagtagcactcgtcaacgtagcctctgtctgcacacagcac 1245
Db 1270 tcaccaatgcatcaaaagtagcactcgtcaacgtagcctctgtctgcacacagcac 1329
Qy 1246 gggggccaagatgcccgctcccgctctg- -----gctggccgggagcctgatcacctgc 1299
Db 1330 gggggccaagatgcccgctcccgctctgctccagctggcgggagcctgatcacctgc 1389
Qy 1300 cctgtgagtcgccaggtgagcctcagctcctcctcctcctcctcctcctcctcctcctc 1359
Db 1390 cctgtgagtcgccaggtgagcctcagctcctcctcctcctcctcctcctcctcctcctc 1449
Qy 1360 aacacacagatttgagctcagcctggtggcagaggtacacacttg- ----- 1408
Db 1450 aacacacagatttgagctcagcctggtggcagaggtacacacttg- ----- 1509
Qy 1409 -----ttgatgatttaataagaaatgt 1428
Db 1510 agtgaggcatcgcaatgtaagactcggtatgatacacactgttggatttaataagaaatgt 1569
Qy 1429 ttacagatccccaagcctggcaaggaaatttctcaactcctgcctcctcctcctcctcct 1488
Db 1570 ttacagatccccaagcctggcaaggaaatttctcaactcctgcctcctcctcctcctcct 1629
Qy 1489 atcaaaagacaccattttggcaagctctatcaccagagcacaacacatcctacaagacac 1548
Db 1630 atcaaaagacaccattttggcaagctctatcaccagagcacaacacatcctacaagacac 1689
Qy 1549 agtgaccataactaattataaccctcgtcaag-ccagcttgaacaccttacttaggaagct 1607

```
|||||
Db 1690 agtgaccataactaattataccccctgcaagcccagcttgaaaccttcacttaggaacgt 1749
Qy 1608 aatcgtgtccccctatacctacttcccccttcccttaattccacagctgctcaataaaagtacaag 1667
Db 1750 aatcgtgtccccctatacctacttcccccttcccttaattccacagctgctcaataaaagtacaag 1809
Qy 1668 agtttaacagtg 1679
Db 1810 agcttaacagtg 1821
|||
```

Search completed: December 7, 2001, 00:19:28
Job time: 6448 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2001, 22:52:55 ; Search time 148.13 Seconds
(without alignments)
6500.775 Million cell updates/sec

Title: US-09-164-862B-4
Perfect score: 1681
Sequence: 1 ctaggtagctggcaccagga.....tacaagatttaacagtgtg 1681

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 596062 seqs, 286424673 residues

Total number of hits satisfying chosen parameters: 1192124

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New:*
1: /cgn2_6/ptodata/2/pna/pna/PCT_NEW_COMB.seq:*
2: /cgn2_6/ptodata/2/pna/pna/US06_NEW_COMB.seq:*
3: /cgn2_6/ptodata/2/pna/pna/US07_NEW_COMB.seq:*
4: /cgn2_6/ptodata/2/pna/pna/US08_NEW_COMB.seq:*
5: /cgn2_6/ptodata/2/pna/pna/US09_NEW_COMB.seq:*
6: /cgn2_6/ptodata/2/pna/pna/US60_NEW_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|------------------------------------------|
| 1 | 389.4 | 23.2 | 414 | 5 | US-09-985-678-249062, Sequence 249062, |
| 2 | 386.6 | 23.0 | 1702 | 5 | US-09-971-392-173, Sequence 173, App |
| 3 | 382.2 | 22.7 | 1496 | 4 | US-08-850-348-1, Sequence 1, Appli |
| 4 | 353.2 | 21.0 | 371 | 5 | US-09-898-888A-36444, Sequence 36444, A |
| 5 | 111.6 | 6.6 | 513 | 5 | US-09-898-888A-28453, Sequence 28453, A |
| 6 | 72.4 | 4.3 | 324 | 5 | US-09-983-965-844, Sequence 844, App |
| 7 | 46.6 | 2.8 | 463 | 5 | US-09-388-906A-19035, Sequence 19035, A |
| 8 | 38 | 2.3 | 352 | 5 | US-09-898-888A-35637, Sequence 35637, A |
| 9 | 36.8 | 2.2 | 545 | 5 | US-09-985-678-212678, Sequence 212678, |
| 10 | 35.6 | 2.1 | 1086 | 5 | US-09-724-797-77, Sequence 77, Appl |
| 11 | 34.8 | 2.1 | 256 | 5 | US-09-983-965-1012, Sequence 1012, Ap |
| 12 | 34.6 | 2.1 | 406 | 5 | US-09-922-340-10241, Sequence 10241, A |
| 13 | 34.6 | 2.1 | 406 | 5 | US-09-922-340A-10241, Sequence 10241, A |
| 14 | 34.6 | 2.1 | 473 | 5 | US-09-922-340-9698, Sequence 9698, Ap |
| 15 | 34.6 | 2.1 | 473 | 5 | US-09-922-340A-9698, Sequence 9698, Ap |
| 16 | 34.6 | 2.1 | 2911 | 5 | US-09-898-888A-1886, Sequence 1886, Ap |
| 17 | 34.4 | 2.0 | 260 | 5 | US-09-985-678-294427, Sequence 294427, |
| 18 | 34.4 | 2.0 | 291 | 5 | US-09-985-678-17720, Sequence 17720, A |
| 19 | 34.4 | 2.0 | 1393 | 5 | US-09-898-888A-14671, Sequence 14671, A |
| 20 | 34.2 | 2.0 | 264 | 5 | US-09-985-678-146271, Sequence 146271, A |
| 21 | 34 | 2.0 | 328 | 5 | US-09-985-678-101133, Sequence 101133, |
| 22 | 34 | 2.0 | 335 | 5 | US-09-849-526A-22977, Sequence 22977, A |
| 23 | 34 | 2.0 | 436 | 5 | US-09-985-678-75847, Sequence 75847, A |
| 24 | 34 | 2.0 | 517 | 5 | US-09-985-678-220810, Sequence 220810, |
| 25 | 34 | 2.0 | 524 | 5 | US-09-985-678-220345, Sequence 220345, |
| 26 | 34 | 2.0 | 536 | 5 | US-09-985-678-219794, Sequence 219794, |
| 27 | 34 | 2.0 | 589 | 5 | US-09-985-678-222994, Sequence 222994, |

ALIGNMENTS

RESULT 1
US-09-985-678-249062
; Sequence 249062, Application US/0985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 16517.255/38-21(15097)F
; CURRENT APPLICATION NUMBER: US/09/985,678
; CURRENT FILING DATE: 2001-11-05
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 249062
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Glycine max
US-09-985-678-249062

| | | | | |
|-----------------------|--------------|-------------------------------------------------------------|---------------|-------------------|
| Query Match | 23.2%; | Score 389.4; | DB 5; | Length 414; |
| Best Local Similarity | 98.3%; | Pred. No. 8.8e-105; | | |
| Matches 404; | Conservative | 0; | Mismatches 6; | Indels 1; |
| Gaps 1; | | | | |
| QY | 8 | gctggcaccaggagcgtggtggaagagagccacacccctgctctgtgagcgc | 67 | Sequence 1015, Ap |
| DB | 4 | gctggcaccaggagcgtggtggaagagagccacacccctgctctgtgagcgc | 63 | Sequence 69, Appl |
| QY | 68 | cagaatgggtggaaggcgcttcaaacaggcttgggtcctgggtgctcagtgctg | 127 | Sequence 23, Appl |
| DB | 64 | gagaatgggtggaaggcgcttcaaacaggcttgggtcctgggtgctcagtgctg | 123 | Sequence 1151, Ap |
| QY | 128 | ctctgcatacaactggctgctgtactacacagctggtcccaagtagcgggaagggatgg | 187 | Sequence 31, Appl |
| DB | 124 | ctctgcatacaactggctgctgtactacacagctggtcccaagtagcgggaagggatgg | 183 | Sequence 1, Appl |
| QY | 188 | gagctgtctccagatgacctgtgacgcttctgtgtaccacatcatcagctttgc | 247 | Sequence 9698, Ap |
| DB | 184 | gagctgtctccagatgacctgtgacgcttctgtgtaccacatcatcagctttgc | 243 | Sequence 9698, Ap |
| QY | 248 | caataagcaacgatcacatcgacacctgggagtggaatgatgtgacgtctacggcat | 307 | Sequence 14344, A |
| DB | 244 | caataagcaacgatcacatcgacacctgggagtggaatgatgtgacgtctacggcat | 303 | Sequence 5908, A |
| QY | 308 | gctcaacacactcaacacacg-aaacccaaactgaagactctctgtgtcgaggat | 366 | Sequence 294, App |
| DB | 304 | gctcaacacactcaacacacg-aaacccaaactgaagactctctgtgtcgaggat | 363 | Sequence 8082, A |
| QY | 367 | ggaactttgggtctcaagattttccaagatagcttccacaccagagtc | 417 | Sequence 9697, Ap |


```
; TELEFAX: 610-270-5090
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 1496 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-08-850-348-1

      22.7%;   Score 382.2;   DB 4;   Length 1496;
Best Local Similarity 59.8%;   Pred. No. 2.3e-102;
Matches 680;   Conservative 0;   Mismatches 448;   Indels 9;   Gaps 2;

QY  43  acaccctgcctcgtcgtcgcagccgaagaatgggtgtgaaggcgctcacaacaggctttg 102
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 103  tggctcgtggtcgtccagtcgtcctcgcacatacaaaactggtcgtcctactacacagct 162
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 146  tgggtcttgcgtctcctcaggagagatcgtcctacaaaactgggttgctactttaccacact 205
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163  ggtccagtagccgggaaggcgatggggagctgctccacagatgccccttgacgcgtctcctgt 222
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 206  ggtccaggaccggcagaccaggaataatccccccttgagataatttgacccttctctat 265
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 223  gtaccacatactacacagcttttgccaaataaagcaacagatcaatcagacacctggaggt 282
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 266  gctctcatctcatctatttcatttgcagcagcatcgaaaacaacaggtttatcaaggaca 325
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 283  ggaatgatgaagctacggcatgctcaacacatcaacacacacacacacacacacacacacga 342
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 326  agagtgaagtgatcgtctaccagaccatcaacagctctcaaaacacgaatccccaaactga 385
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 343  agactctctgtcgtcgaggatgggaactttgggtctcaaaagattttccaagatagcct 402
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 386  aaattcttctgtccattggagggttacctgttgggtccaaagggttccaccctatgtgtg 445
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 403  ccaacacccagagtcgcgcgacttctcaagtcagtcaccgcaatttctgcgacccatg 462
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 446  attcttctacatcacgcttggaatttcaatttaactccataatcctgtttctgaggaaaccata 505
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 463  gctttgatggcgtagccttgcctgctctaccctgacgcggagagacaaacacacatttta 522
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 506  actttgatgacttggatgttaagcttggatctaccacagatcagaaaagaaaacacactatttca 565
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523  ccaccctaatcaaggaaatgaagccgaatttataaagggaagcccgaggggaaagc 582
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 566  ctgtgcttgattcatgatttagcgaagcccttttcagaaggacttcacaaaatccaccagg 625
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 583  a---gctcctgctcagcgcagcactgctcgtcggggaagggtcacatttgacagcagctatg 639
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 626  aaaggcttcttctgacttcggtgatactgcaggagggcaaaatgattgataacagctatc 685
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640  acattgccaagatatcccaacacctgatttcattagcatactgacatcagatttcatg 699
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 686  aagtttgagaaactggcaaaaagatctggaatttcattcaacacctcctgtcctttgacttccatg 745
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700  gcgcctggcgtggagc-----cacaggccatcacagtcctccctcagcgaggtcaggg 753
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 746  ggtcttgggaaagcccttattactgcccacacacaccccttctgaggcaagggtggcagg 805
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 754  atgcaagctctgacagattcagcaacactgactatgctgtgggtacatgtttgaggtgg 813
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 806  acagaggcccaagctctacttacaataatgttggaattatgtgtgggtacttggaatacataagg 865
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 814  gggcctcgtccagtaagctgggtgaggtccccacacctcgtggagagagcttcaactctgg 873
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 866  gaattgccatcagagaaggtggtcatgggcatccccacatatatgggcactcctcttccactgg 925
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 4
US-09-898-888A-36444
; Sequence 36444, Application US/09898888A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; TITLE OF INVENTION: LIBRARIES
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/898,888A
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/340,623
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/205,070
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 36444
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(371)
; OTHER INFORMATION: n = A,T,C or G
US-09-898-888A-36444

Query Match 21.0%; Score 353.2; DB 5; Length 371;
Best Local Similarity 98.7%; Pred. No. 3.8e-94;
Matches 366; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 34  gaagagccacacacctgctcgtcgtcagcagaatgggtgtgaaggcgtctctcaaa 93
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 94  caggcttggctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 153
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61  caggcttggctcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgtcgt 120
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 154  acaccagctggtccagtcaccgggaagcgatggagctgcttcccagatgccccttgacc 213
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121  acaccagctggtccagtcaccgggaagcgatggagctgcttcccagatgccccttgacc 180
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 214  gcttcctgtgtaccacacatcatctacagctttgtccaatataagaacagatcacatcgaca 273
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181  gcttcctgtgtaccacacatcatctacagctttgtccaatataagaacagatcacatcgaca 240
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 274  cctggagatggaatgatgtacgctctacggcatgctcaacacactcaacacacagcaacc 333
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241  cctggagatggaatgatgtacgctctacggcatgctcaacacactcaacacacagcaacc 300
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 334  ccaacctgaagaactctctgtcgtcggaggtggaactttggg-tctcaagaagattttcc 392
Db    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 301 ccaactgaagactcttctgtctgcgaggatggaacttgggnctcaagattttcc 360
QY 393 aagatagcctc 403
Db 361 aagatagcctc 371

RESULT 5
US-09-898-888A-28453
; Sequence 28453, Application US/09898888A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; LIBRARIES
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/898,888A
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/340,623
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US 09/205,070
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 28453
; LENGTH: 513
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(513)
; OTHER INFORMATION: n = A,T,C or G
US-09-898-888A-28453

Query Match 6.6%; Score 111.6; DB 5; Length 513;
Best Local Similarity 53.1%; Pred. No. 6.5e-23;
Matches 263; Conservative 0; Mismatches 223; Indels 9; Gaps 2;

QY 335 caactgaagactcttctgtctgcgaggatggaacttgggtctcaagattttccaa 394
Db 2 caaatgaattcttctgtccataggggtgaactgttgggttccaaagggttccacc 61
QY 395 gatagcctccacacccagagtcgcggactttccataagtcagtcagcgcattttgcg 454
Db 62 tatgtgatctttacatcacgcgttggaattcatttaactccataatcctgtttctgag 121
QY 455 caccatggtttgatggcgtgaccttgctgctgctctaccctgagcggagagacaaca 514
Db 122 gaaccataactttgatggactggatgtaagctggatccaccagatcagaagaacacac 181
QY 515 ccatittaccaccttaatacaaggaaatgaaggccgaatttataagggaagccagccagg 574
Db 182 tcatttcactgtgattcatgatagtagcagaagccttcagaaggacttcacaaaac 241
QY 575 gaaaaagca---gctcctgctcagcgcagcactgtctgcggggaaggctcaccattgacag 631
Db 242 caccagaagaaggctctcttgactgcgggcgtatctcaggggcgaatgattgataa 301
QY 632 cagctatgacatgccaagatcccaacacactgatttattcattagcatcatgacacacga 691
Db 302 cagctatcaagttagaactggcaaaagatctggtatttcataaccctctgctcttga 361
QY 692 ttttcatgcgcctgg-----cgtgggaccacagccatcacagtcctccctcaggcagg 745
Db 362 ctctccatgggtctgggaaaaaaccccttactactgcccacacagccctctgagcagag 421
QY 746 tcaggagagatgaagtcctgcagacagattcagcacaactgactatgctgtgggtacatgtt 805
Db 422 gtggcagacacagagncnagctnccnanaannnngggnaaancnnggggncnnggga 481
QY 806 gaggtctgggggtctc 820
|||

Db 482 nanatngggganncc 496

RESULT 6
US-09-983-965-844
; Sequence 844, Application US/09983965
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Tyatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 844
; LENGTH: 324
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 30-LIB188-016-Q1-E1-H9
US-09-983-965-844

Query Match 4.3%; Score 72.4; DB 5; Length 324;
Best Local Similarity 51.9%; Pred. No. 1.8e-11;
Matches 163; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

QY 610 cgggggaaggtccaccatgacagcagctatgacattgccaagatatcccaacactgatt 669
Db 6 cagggaacactaccattaccagaggccaagcaatcccaactaagtgcacacctgggcc 65
QY 670 tcattagcatcatgactcagattttcattgctgcctggcgtgggaccacagggccatcaca 729
Db 66 tcacaggcatcttcacgcacgtgctcagacagatccctgtaccacaacagtcgcgaacct 125
QY 730 gtccctccaggcagggtcaggaggtgcaagtcctcagacagattcagcaacactgactatg 789
Db 126 acgacgggtatgatgtcactacttccactagttcagttgcctacacagattaccaactgt 185
QY 790 ctgtggggtacatgtttagcgtggggctcctcctcagtaagctggtgagtggtggtatcccca 849
Db 186 cagttcgtgacagctttacgcgggtgggtcccaacataagacggtttatgagtaacacta 245
QY 850 ccttcgggaggagcttcactctgtcttctcttgagactggtgttcacagcgccaactcag 909
Db 246 cctttgcatcataactcgtcagcggcatccatttgatgtgtgagaccacatgtcaa 305
QY 910 gaccgggaattcca 923
Db 306 tgacactaatcca 319

RESULT 7
US-09-388-906A-19035
; Sequence 19035, Application US/09388906A
; GENERAL INFORMATION:
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Shenk, Michael Andrew
; TITLE OF INVENTION: Polynucleotides Isolated From Plants and
; TITLE OF INVENTION: Methods For Their Use
; FILE REFERENCE: 11000.10130
; CURRENT APPLICATION NUMBER: US/09/388,906A
; CURRENT FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 24843
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19035
```

```
; LENGTH: 463
; TYPE: DNA
; ORGANISM: Eucaalyptus grandis
US-09-388-906A-19035

Query Match
Best Local Similarity 2.8%; Score 46.6; DB 5; Length 463;
Matches 112; Conservative 0; Mismatches 109; Indels 0; Gaps 0;

QY 268 tcgacacctggagtggaatgatgtgacgctctacgcatcgtcacacactcaacaaca 327
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 65 tgaagctcggactcgacagccccgcgtctcgaactcactgcctccctccgcgga 124

QY 328 cgaaccccaacctgaagactctctgtctgtcgaggagtggaactttgggtctcaaat 387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 125 agaaccccgccgtgaagacccttatccatcgcgcgaggagtgatggccctactctgt 184

QY 388 ttccaagatagctcccaacacccagagtcgcccgaactttcatcaagtacgtaccgccaat 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 185 ttgcacacattgcacgtgcttcgacgaggaagctctctcatagactcctccattgagg 244

QY 448 ttctgcacccatgcttggatggcgctgaccttgctgg 488
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 245 tcgcgaggaagtacagtgttggatggctcgatctcgattgg 285

RESULT 8
US-09-898-888A-35637/c
; Sequence 35637, Application US/0989888A
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc.
; TITLE OF INVENTION: NOVEL CONTIGS OBTAINED FROM VARIOUS CDNA
; FILE REFERENCE: 20411-748CON1
; CURRENT APPLICATION NUMBER: US/09/898.888A
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/340.623
; PRIOR FILING DATE: 1999-06-28
; PRIOR APPLICATION NUMBER: US/09/205.070
; PRIOR FILING DATE: 1998-12-03
; NUMBER OF SEQ ID NOS: 45207
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35637
; LENGTH: 352
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-898-888A-35637

Query Match
Best Local Similarity 2.3%; Score 38; DB 5; Length 352;
Matches 53; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 1093 tgaagataggcagctgcgagcgccatgtatggccctggacctggatgacttccagg 1152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 TAAAGATTAACTGGGAGGAGCCATGCTGTCTATTGACATGGATGACTTCAGTG 292

QY 1153 gctccttctgcggccagg 1170
    || ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 291 GCAAAATCCTGCAACCAGG 274

RESULT 9
US-09-985-678-212678
; Sequence 212678, Application US/09985678
; GENERAL INFORMATION:
; APPLICANT: Cheikh, Nordine
; APPLICANT: Liu, Jingdong
; TITLE OF INVENTION: Annotated Plant Genes
; FILE REFERENCE: 16517.255/38-21(15097)F
; CURRENT APPLICATION NUMBER: US/09/985.678
; CURRENT FILING DATE: 2001-11-05
```

```
; PRIOR APPLICATION NUMBER: US 09/304,517
; PRIOR FILING DATE: 1999-05-06
; NUMBER OF SEQ ID NOS: 295529
; SEQ ID NO 212678
; LENGTH: 545
; TYPE: DNA
; ORGANISM: Glycine max
US-09-985-678-212678

Query Match
Best Local Similarity 2.2%; Score 36.8; DB 5; Length 545;
Matches 125; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 169 agtacccggaaagcgatgggagctgtctccagatgcccttgaccgcttctctgttaccc 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 128 agcccaaggccgcgaacaccccgacccctcaaaacctccaaatctaccggtggaacc 187

QY 229 acatcatctacagcttgcgaataaagcaacgatacacatcgacacctgggagtggaatg 288
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 188 ccgaaacccctcaagcccgagctcaaggactaccagatacaactaaaggagtgaggc 247

QY 289 atgtgacgctctacgcatgctcaacacactcaacaacacacacacacacacacacacac 348
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 ccattgctctcgacgctcatcaagaatcaagaacgagatcgaccgagcctcacgttcc 307

QY 349 tcttctgtcggaggatggaaactttggctctcaagattttccaaagatagcctccaaca 408
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 308 gccgctcgtgcgcgagggtatctgcgtcctcgcgaatgaacatcgacggtgcaacg 367

QY 409 ccagagtcgcccgaactttcatcaagtcagta 440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 368 gctcgtgctgcctcccaagaatccctcggaa 399

RESULT 10
US-09-724-797-77/c
; Sequence 77, Application US/09724797
; GENERAL INFORMATION:
; APPLICANT: Jon S. THORSON
; TITLE OF INVENTION: MICROMONOSPORA ECHINOSPORA GENES
; TITLE OF INVENTION: ENCODING FOR BIOSYNTHESIS OF
; TITLE OF INVENTION: CALICHEAMICIN AND SELF-RESISTANCE THERETO
; FILE REFERENCE: 2653-40
; CURRENT APPLICATION NUMBER: US/09/724,797
; CURRENT FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 60/111,325
; PRIOR FILING DATE: 1998-12-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 77
; LENGTH: 1086
; TYPE: DNA
; ORGANISM: Bacteria
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1086)
US-09-724-797-77

Query Match
Best Local Similarity 2.1%; Score 35.6; DB 5; Length 1086;
Matches 113; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 1103 gcagctggcaggcgccatggtatggcccttgacctggacttccagggtctctctg 1162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 712 GCAGCGGGCGGTGGCCCGCCCGGCGACCCCGCACCGGACGTCCTCTCTCGGCGACG 653

QY 1163 cggccaggatcgtcgttctctctcaccatgccatcaaggatgcactcgtgcaacgta 1222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 CGGTACCCGTCACCGTTCGGTGGTTCATCAGGGCGGCGTGCAGCGACTCGGCGACGCA 593

QY 1223 gccctctgttgcacacagcagggggcgaaggatgccccgtccccgtctgctggcgg 1282
```

```

; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(406)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-10241

Query Match      2.1%; Score 34.6; DB 5; Length 406;
Best Local Similarity 53.3%; Pred. No. 2.7;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 320 caacaacagaaaccccaacctgaagactcttctgtcgcgagagatggaactttgggtc 379
    || || || || || || || || || || || || || || || || || || || || ||
DB 221 catcaggaagtagtacaagtagaagtcttctcatcagtcacaaatggatcttggctc 280
    || || || || || || || || || || || || || || || || || || || || ||

QY 380 tcaagattttccaagatagctcctcaacacccagagtcgcccgaactttcatcaagtca 439
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 281 tgttagcgtctccaccgtggcctcattccttgtgaggtcaatgaccttgggttgcct 340
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 440 accgccatttctgcga 456
    ||||| || || ||
DB 341 catgccaattttctgca 357
    ||||| || || ||

RESULT 13
US-09-922-340A-10241
; Sequence 10241, Application US/09922340A
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340A
; CURRENT FILING DATE: 2001-08-03
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/217,517
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10241
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(406)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340A-10241

Query Match      2.1%; Score 34.6; DB 5; Length 406;
Best Local Similarity 53.3%; Pred. No. 2.7;
Matches 73; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 320 caacaacagaaaccccaacctgaagactcttctgtcgcgagagatggaactttgggtc 379
    || || || || || || || || || || || || || || || || || || || || ||
DB 221 catcaggaagtagtacaagtagaagtcttctcatcagtcacaaatggatcttggctc 280
    || || || || || || || || || || || || || || || || || || || || ||

QY 380 tcaagattttccaagatagctcctcaacacccagagtcgcccgaactttcatcaagtca 439
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 281 tgttagcgtctccaccgtggcctcattccttgtgaggtcaatgaccttgggttgcct 340
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 440 accgccatttctgcga 456
    ||||| || || ||
DB 341 catgccaattttctgca 357
    ||||| || || ||

US-09-983-965-1012
; Sequence 1012, Application US/09983965
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 37-21(10297)C
; CURRENT APPLICATION NUMBER: US/09/983,965
; CURRENT FILING DATE: 2001-10-26
; PRIOR APPLICATION NUMBER: US 09/465,231
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,678
; PRIOR FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 5912
; SEQ ID NO 1012
; LENGTH: 256
; TYPE: DNA
; ORGANISM: Bos taurus
; FEATURE:
; OTHER INFORMATION: Clone ID: 50-LIB188-021-Q1-E1-E6
US-09-983-965-1012

Query Match      2.1%; Score 34.8; DB 5; Length 256;
Best Local Similarity 52.0%; Pred. No. 1.8;
Matches 78; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 1282 gggagcctgatccactgccctgctgagtcgccagctgagctcagctccctccctctggg 1341
    || || || || || || || || || || || || || || || || || || || || ||
DB 83 ggaagacaagtcctcgtgctgctgagtcacccgggagcttcacgcgcgggctgggc 142
    || || || || || || || || || || || || || || || || || || || || ||

QY 1342 gcctatgcagaggtccacaacacacagagatttgagctcagccctggtgggagagaggtac 1401
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 143 cacaacgcagagccccaacacacacagcagcgtgttcaaaaaggttatgagcgtggaacc 202
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1402 acacttgttgatgattaatgaaatgttta 1431
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 203 atccgtgatgcggtgggattttacagctta 232
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
US-09-922-340-10241
; Sequence 10241, Application US/09922340
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 20411-726CON2
; CURRENT APPLICATION NUMBER: US/09/922,340
; CURRENT FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 09/353,690
; PRIOR FILING DATE: 1999-07-14
; PRIOR APPLICATION NUMBER: US 09/004,182
; PRIOR FILING DATE: 1998-01-07
; NUMBER OF SEQ ID NOS: 12181
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10241
; LENGTH: 406
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(406)
; OTHER INFORMATION: n = A,T,C or G
US-09-922-340-10241
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 6, 2001, 20:19:00 ; Search time 1546.1 Seconds
(without alignments)
11683.367 Million cell updates/sec

Title: US-09-164-862B-4

Perfect score: 1681

Sequence: 1 ctaggtagctggcaccagga.....tacaagagtttaacagtgtg 1681

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 11351937 seqs, 5372889281 residues

Total number of hits satisfying chosen parameters: 22703874

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST: *
1: em_estfun: *
2: em_esthum: *
3: em_estin: *
4: em_estom: *
5: em_estpl: *
6: em_estba: *
7: em_estro: *
8: em_estov: *
9: em_hctc: *
10: gb_estl: *
11: gb_est2: *
12: gb_hctc: *
13: gb_gss: *
14: em_gss_fun: *
15: em_gss_hum: *
16: em_gss_inv: *
17: em_gss_pln: *
18: em_gss_pro: *
19: em_gss_rod: *
20: em_gss_vrt: *
21: em_gss_other: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 864.2 | 51.4 | 971 | 11 | BG675109 602621376 |
| 2 | 791.8 | 47.1 | 924 | 11 | BG742515 602632364 |
| 3 | 787.8 | 46.9 | 928 | 11 | BG743108 602634245 |
| 4 | 770.4 | 45.8 | 821 | 11 | BG743506 602635536 |
| 5 | 766.2 | 45.6 | 830 | 11 | BG743515 602635545 |
| 6 | 762.8 | 45.4 | 892 | 11 | BG741628 602633472 |
| 7 | 752.4 | 44.8 | 885 | 11 | BF338858 602036114 |
| 8 | 746.8 | 44.4 | 907 | 11 | BG696335 602659494 |
| 9 | 746.2 | 44.4 | 882 | 11 | BG740681 602631527 |
| 10 | 733.8 | 43.7 | 976 | 11 | BG697544 602660764 |
| 11 | 722.8 | 43.0 | 810 | 11 | BG696363 602658744 |
| 12 | 722.2 | 43.0 | 1060 | 11 | BG698930 602703349 |

| | | | | | |
|----|-------|------|------|----|--------------------|
| 13 | 711.4 | 42.3 | 91.2 | 11 | BF337376 602035091 |
| 14 | 706 | 42.0 | 81.9 | 11 | BF339902 602036420 |
| 15 | 692.8 | 41.2 | 906 | 11 | BF339615 602035090 |
| 16 | 682.4 | 41.2 | 908 | 11 | BG740404 602634181 |
| 17 | 689.2 | 41.0 | 1054 | 11 | BF525672 602069793 |
| 18 | 687.6 | 40.9 | 722 | 11 | BG739851 602630581 |
| 19 | 679.4 | 40.4 | 1091 | 11 | BF343234 602015887 |
| 20 | 676.8 | 40.3 | 1137 | 11 | BF338905 602036163 |
| 21 | 671.8 | 40.0 | 1054 | 11 | BF342106 602012863 |
| 22 | 670.4 | 39.9 | 865 | 11 | BF525695 602070024 |
| 23 | 660.8 | 39.3 | 1063 | 11 | BF340459 602037080 |
| 24 | 656 | 39.0 | 703 | 11 | BF340959 602038274 |
| 25 | 646 | 38.4 | 1134 | 11 | BF525348 602069425 |
| 26 | 645.2 | 38.4 | 830 | 11 | BF341104 602038121 |
| 27 | 641.2 | 38.1 | 969 | 11 | BG697461 602660650 |
| 28 | 637 | 37.9 | 729 | 10 | BE439671 HTM1-341R |
| 29 | 629.6 | 37.5 | 969 | 11 | BF340511 602037143 |
| 30 | 629.2 | 37.4 | 883 | 11 | BG697609 602660850 |
| 31 | 628.2 | 37.4 | 886 | 11 | BF525836 602069987 |
| 32 | 627.4 | 37.3 | 1026 | 11 | BG698054 602659909 |
| 33 | 618.4 | 36.8 | 671 | 11 | BF348690 602034410 |
| 34 | 613.4 | 36.5 | 646 | 11 | BF343345 602017516 |
| 35 | 613 | 36.5 | 834 | 11 | BF338293 602035520 |
| 36 | 612 | 36.4 | 1017 | 11 | BG743545 602635583 |
| 37 | 610.4 | 36.3 | 827 | 11 | BF339047 602034840 |
| 38 | 607.4 | 36.1 | 644 | 10 | BE439481 HTM1-111F |
| 39 | 604.2 | 35.9 | 730 | 11 | BF343573 602014404 |
| 40 | 603.2 | 35.9 | 649 | 10 | AI903381 RC-HT029- |
| 41 | 602 | 35.8 | 671 | 11 | BF339330 602038660 |
| 42 | 600.6 | 35.7 | 710 | 11 | BF344188 602017434 |
| 43 | 595.6 | 35.4 | 637 | 11 | BF341141 602038173 |
| 44 | 572.8 | 34.1 | 754 | 11 | BF340255 602036845 |
| 45 | 565.4 | 33.6 | 739 | 11 | BF525705 602070035 |

ALIGNMENTS

RESULT 1

BG675109 602621376F1 NCI_CGAP_Skn3 Homo sapiens EST 01-MAY-2001
LOCUS mRNA sequence.
DEFINITION 971 bp mRNA
ACCESSION BG675109
VERSION BG675109.1 GI:13906505
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 971)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10595 row: h column: 19
High quality sequence stop: 895.

FEATURES

source
Location/Qualifiers
1..971
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4746690"
/clone_lib="NCI_CGAP_Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;


```
365 atggaacttgggtctcaagaattttccaaagtagctccaaacacccagagtcgcggac 424
Db 361 ATGGAACCTTTGGGTCTCAAAAGATTTTCCAAAGATAGCTTCCAAACACCCAGAGTCGCCGGAC 420
QY 425 ttgcatcaagtcagtcaccgcccatttctgcgaccccatggttgcgttgatggcgtagccttgc 484
Db 421 TTTCAATCAAGTCAGTACCGCCCATTTCTGGCGACCCCATGCTTTGATGGGTGGACCTTGC 480
QY 485 ctggctctaccctggagcagagagacaacaccattttaccacccctaatcaagaaatgaa 544
Db 481 CTGGCTCTACCTTGGACGGAGAGACAACAGCATTTTACCACCCCTAATCAAGGAATGAA 540
QY 545 ggcgaattataaagaagccagccagcagggaaagcagctctctcagcgagcact 604
Db 541 GGCCGAATTTATAAGGAAGCCAGCCA -GGAAAAGCAGCTCTCTGCTCAGCGCAGCACT 599
QY 605 gtctcg99gaaggtccaccattgacgcgctatgacattgccaagatgcccaacacct 664
Db 600 GTCTCGGGGAAGGTCAACATTGACAGCAGCTATGACATTGCCAAGATATCCCAACACCT 659
QY 665 ggaattcattagcatcatgacactaga-ttttcattgagcgcctggcgtggaccacaggcc 723
Db 660 GGATTTCTATTAGCATCATGACCTTACCATTTTTTCATGGAGCCTGGCGTGGACACAGGCC 719
QY 724 atcacagtcctccctcaggcga-ggtcagaggatgcgaagtccttcagagattcagcaaacact 782
Db 720 ATCACAGTCCCTTGTCCAGCGGTCAAGAGGATGCAAGTCTCTGACAGATTCAGCAACACT 779
QY 783 gactatgctgtg99gtacatgttgtaggctgg99ggtccctgcagtagtaagctgggtgagggc 842
Db 780 GACTATGCTGTGGGTACATGTTGACGGGTGGGGCTCTCCAGTAAGCTGGTGTGATGGGC 839
QY 843 atcccacacttcg99gagggctccactcgtgctctctcttcgagactggtgtccagcgcca 902
Db 840 TCCACCTCTCGGGAGGAAGTTCACCTTCTGGTCTTTCTGACACCTGGTGTGGGCCCAT 899
QY 903 atc 905
Db 900 CTC 902

RESULT 3
BG743108 928 bp mRNA EST 15-MAY-2001
LOCUS 602634245F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4779371 5',
DEFINITION mRNA sequence.
ACCESSION BG743108
VERSION BG743108.1 GI:14053761
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 928)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/,
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
plate: LLAM10636 row: j column: 12
High quality sequence stop: 845.
FEATURES
Location/Qualifiers
source
1..928
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4779371"
```

```
/clone_lib="NCI_CGAP_Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/Note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 227 a 256 c 249 g 196 t
ORIGIN
Query Match 46.9%; Score 787.8; DB 11; Length 928;
Best Local Similarity 94.9%; Pred. No. 8.3e-190;
Matches 846; Conservative 0; Mismatches 42; Indels 3; Gaps 3;
QY 1 ctaggtagctggcaccagagcggtggcaagggaagagggccaacacccctgacctctg 60
Db 33 CTAGGTAGCTGGCACCAGGAGCGGTGGCAAGGAAGAGGCCACACCTGCCCTGCTCTG 92
QY 61 ctgcagcccaaatgggtgtgaaggcgtctcaaacagggcttggctcctgggtgctctcc 120
Db 93 CTGCAGCCAGATGGGTGTGAAGCGCTCTCAACAGAGCTTTGTGGTCTGCTGCTCC 152
QY 121 agtctgctctgcatacaaaactggctctgctactacaccagctggtccccagtagccggaag 180
Db 153 AGTGCTGCTCTGCATACAAACTGGTCTGTCTACTACACAGCTGCTCCAGTACCGGAAG 212
QY 181 gqatgggagctgtccccagatgccccttgacccgtctcctgtgtaccacatcatctaca 240
Db 213 GCGATGGGAGCTGCTTCCAGATGCCCCCTTGACCGCTTCTCTTACCCACATCATACA 272
QY 241 gctttgccataaagaacagatcacatcgacacctggagtggaatgatgtgacctct 300
Db 273 GCTTTGCCAATATAAGCAAGATCATATGCACACCTTGGGAGTGGAAATGATGACGCTCT 332
QY 301 acggctgctcacaacactcaacaacacgaaacccccacccacccacccacccacccaccc 360
Db 333 ACGGATGCTCAACACACTCAAGAACAGGAAACCCCAACCTGGAAGACTCTCTTCTGTCTG 392
QY 361 gaggatggaacttgggtctcaaaagattttccaaagatagcctcccaacacccagagtcgc 420
Db 393 GAGGATGGAACCTTTGGGTCTCAAAAGATTTTCCAAAGATAGCTCTCCAAACACCCAGAGTCGCC 452
QY 421 ggaatttcacatcagtcagtcaccgcccattctgcgacccatgcttgcgtggcgtagcc 480
Db 453 GGACTTTCAATCAAGTCAGTACCCCATTTCTGCGACCCCATGCTTTGTATGGCTGGACC 512
QY 481 ttgcctggctctaccctggagcagagagacaacacacattttaccacccctaatcaagaaa 540
Db 513 TTGCTGGCTCTTACCCTGGAGGGAGAGACAACACAGCATTTTACCACCCCTAATCAAGGAAA 572
QY 541 tgaaggccgaattataaagggaagccagccaggggaaaaaagcagctcctgctcagcgacg 600
Db 573 TGAAGCCCGAATTTAATAAGGAAGCCAGCCAGGAAAAAGCAGCTCTCTCTCAGCGCAG 632
QY 601 cactgctcggggaaggtccaccattgacagcagctatgacattgccaagatcccaaac 660
Db 633 CACTGTCTCGGGGAAGGTCAACCTTTGACAGCAGCTATGACATTTGCCAAGATATCCCAAC 692
QY 661 acctgatttcattagcatcatgacacatgattttcatggcgcctggcggtgggaccacag 720
Db 693 ACCTGGATTTTCAATACATCATCACTACGATTTTTCATGGAGCCTGGCGTGGGA-CACAG 751
QY 721 gccatcacagtcctccctcaggcgaggttcaggagatgcaagtcctctgacagattcagaaca 780
Db 752 GCCATCACAGT-CCCTGGTCCGAGGTCAAGGAGATGCAAGTCTCTGACAGATTTCAGCAACA 810
QY 781 ctgactatgctgggtgtacatgttaa-ggctgggggctcctccagtagagctgggtagt 839
Db 811 CTGACTATGCTGTGGCGTACATGTTGACGGCTGGGGGCTCCTTCAATAAAGCTGTGAATG 870
QY 840 ggcacccacaccttcgggagagagcttcaactctggtctctctctgagagctggt 890
Db 871 GGCATCCCAACCTTTCGGAAGGAGCTCTACTCGGGTTCITTCGAACCGGGCT 921
```

| | | | |
|-----------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------|
| Db | 361 | ATGGAACCTTGGGTCTCAAGATTTTCCAAAGATAGCTTCCAAACACCCAGAGTCGCGGAC | 420 |
| QY | 425 | tttcatcaagtcagtaacgcatttctgcgaccatggtttgatggcgtgaccttgc | 484 |
| Db | 421 | TTTCATCAAGTCAGTACGCCATTTCTGCGCACCATGGCTTTGATGGCTGGACCTTGC | 480 |
| QY | 485 | ctggctctaccctgacgagagacaaacacattttaccacccttaatacaaggaatgaa | 544 |
| Db | 481 | CTGGCTCTACCCCTGGACGGAGAGACAACACAGCATTTTACCACCTTAATCAAGGAATGAA | 540 |
| QY | 545 | ggccgaattataaaggaagccagccaggggaaagcagctcctgcctgcagcgcagcact | 604 |
| Db | 541 | GGCGAATTTATAAGGAAGCCAGCCAGGGAAAAGAGCTCTCTGCTCAGCGCAGCACT | 600 |
| QY | 605 | gtctgagggaagggtccaccattgacagcagctatgacattgccaagatatcccaacact | 664 |
| Db | 601 | GTCTGGGGGAAGGTCAACATTCACAGCAGCTATGACATTGCCAAGATATCCCAACACT | 660 |
| QY | 665 | ggatttcattagcatcatcaccctacgattttcatggcgcctgctgggaccacagggcca | 724 |
| Db | 661 | GGATTTCATTAGCATCATGACCTACCATTTTCTATGGAGCCTGGCGTGGCCACAGGCCA | 720 |
| QY | 725 | tcacagtcctccctcagcgaggt-caggagatgcagagtcctgcagcattcagcaacactg | 783 |
| Db | 721 | TCACAGTCCCTCTTCGAGGTCAGAGGATGCAAGTCTGACAGATTTCAGCAACACTG | 780 |
| QY | 784 | actatgctgggggta--catgttggagctgggggctcct | 821 |
| Db | 781 | ACTATGCTGTGGGGTAACATGTTTGGAGCTGGGGGCTCCT | 820 |
| RESULT | 5 | | |
| LOCUS | 602635545F1 | NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4780522 | 5' |
| DEFINITION | 830 bp | mRNA | EST 15-MAY-2001 |
| ACCESSION | BG743515 | mRNA sequence. | |
| VERSION | BG743515.1 | GI:14054168 | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | 1 | (bases 1 to 830) | |
| AUTHORS | NIH-MGC | http://mgc.nci.nih.gov/ | |
| TITLE | National Institutes of Health, Mammalian Gene Collection (MGC) | | |
| JOURNAL | Unpublished (1999) | | |
| COMMENT | Contact: Robert Strausberg, Ph.D. Email: cgapbs-r@mail.nih.gov Tissue Procurement: James Cleaver, M.D. cDNA Library Prepared by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLAM10639 row: j column: 11 High quality sequence stop: 826. Location/Qualifiers | | |
| FEATURES | source | | |
| | 1..830 | /organism="Homo sapiens" | |
| | | /db_xref="taxon:9606" | |
| | | /clone="IMAGE:4780522" | |
| | | /lab_host="DH10B (T1 phage-resistant)" | |
| | | /note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI_CGAP Library." | |
| BASE COUNT | 209 a | 232 c | 220 g |
| ORIGIN | 169 t | | |
| Query Match | 45.8% | Score 770; | DB 11; Length 821; |
| Best Local Similarity | 97.8% | Pred. No. 2.7e-185; | |
| Matches | 802; | Conservative 0; | Mismatches 15; Indels 3; Gaps 2; |
| QY | 5 | gtagctggcaccagagcgtggtgggcaagggaagagccacacccctgctctgctgc | 64 |
| Db | 1 | GTAGCTGGCACCAGGAGCGGTGGGCAAGGGAAGAGGCCACACCTGCTCTGCTGC | 60 |
| QY | 65 | agccagaatggtgtgaagcctcacaagcgtttgtccctggctgctcagtg | 124 |
| Db | 61 | AGCCAGAATGGGTGTGAAGGCGCTCAACACAGCGCTTGTGGTCCCTGCTCCAGTG | 120 |
| QY | 125 | ctgctctgcatacaaaactggtctgctactacaccagctggtccagtagtaccgggaagcgga | 184 |
| Db | 121 | CTGCTCTGCATCAAACTGCTGCTACTACACAGCTGCTGCCAGTACCGGGAAGCGGA | 180 |
| QY | 185 | tgggagctgtctccagatgcctctgacgccttctgtgtacccacatactacagctt | 244 |
| Db | 181 | TGGGAGCTGCTTCCAGATGCCCTTGACCGCTTCTCTGTATCCACATCATACAGCTT | 240 |
| QY | 245 | tgcataataagcaagatacacatgcagacacctggagtgatgatgacgtctacgg | 304 |
| Db | 241 | TGCAATATAAGCAAGCATCATACACACCTGGGAGTGGATGATGATGACGCTTACGG | 300 |
| QY | 305 | catgctcaacactcaacaacacgaccccaacctgaagactctcttctgtcgtggagg | 364 |
| Db | 301 | CATGCTCAACACTCAAGACAGGAACCCCACTGAAGACTCTCTTCTGTCTCGGAGG | 360 |
| QY | 365 | aaacttgggtctcaagattttccaagatagcgtcccaacacccagagtcgcggac | 424 |

```
Query Match      45.6%; Score 766.2; DB 11; Length 830;
Best Local Similarity 97.5%; Pred. No. 2.5e-184;
Matches 810; Conservative 0; Mismatches 18; Indels 3; Gaps 3;

QY 257 caacgatcacatgcacacctggagtggaatgatgtgacgtctacggtacgtcacaac 316
DB 1 CAACGATCACATCGACACCTGGAGTGGGAATGATGTGACGCTCTACGGCATGCTCAACAC 60

QY 317 actcaacaacagaaccccaa-cctgaagactctctgtctcgagagatgaactttg 375
DB 61 ACTCAAGAACAGGACCAACCAAGCCTGAAGACTCTTGTCTGCGAGGATGAACCTTGG 120

QY 376 ggtctcaagattttccaagatagcttccaacaccagagtcgcccgaactttcatcaagt 435
DB 121 GGTCTCAAGATTTCACAGATAGCTTCCAACACCAGAGTCGCCGACTTTCATCAAGT 180

QY 436 cagtaccgcatttctgcgacccatggctttgatggcggtgacctgacctctacc 495
DB 181 CAGTACCGCATTTCTGCGCACCATGGCTTGTGATGGCTGGACCTTGGCTGCTTACC 240

QY 496 ctggacggagagacaacacattttaccacccctaatacgaagaaatgaagggccgaattta 555
DB 241 CTGGACGGAGAGACAACAGCATTTTACCACCCCTAATCAAGGAATGAAGGCCGAATTA 300

QY 556 taagaagaacccagcagggaaagacagctcctctcagcgagcactgtctgcgggga 615
DB 301 TAAGGAAGCCCAAGCCAGGAAAGACGCTCTGCTCAGCGCAGGACACTGTCTGCGGGGA 360

QY 616 aggtcaccattgacagcagctatgacattgcccgaagatatcccaacacctggatttcatta 675
DB 361 AGGTCAACATTGACAGCAGCATGACATTGGCAAGATATCCCAACACCTGGATTTTATTA 420

QY 676 gaatcagacctacgattttcattgacgctggtggacccacagggccatcacagtcgcc 735
DB 421 GCATCATGACCTACGATTTTTCATGAGCCCTGGCGTGGGACACAGGCCCATTCACAGTCCC 480

QY 736 tcaggcgagtcagggagtgcaagtcctcgacagattcagcaacactgactactgctgtgg 795
DB 481 TGTTCGAGGTGAGGAGGATGCAAGTCTCTGACAGATTACAGCAACTGACTATGCTGTGG 540

QY 796 ggtacattgttgagctggggctccctgccagtaagctgggtgatgggcatccccaccttcg 855
DB 541 GGTACATGTTGAGGCTGGGGCTCCTGCCAGTAAGCTGGTGTGAGTGGGCATCCCCACCTTCG 600

QY 856 gggagggttcactctggtcttcttgagactggtgttccagcgccaatctcagggaccgg 915
DB 601 GGAGGAGCTTCACTCTGGCTTCTCTGAGACTTGTGTGGAGCCCAATCTCAGGACCGG 660

QY 916 gaattccaggcggttccacaaggaagggagggacccttgcctactatgagatctgtgact 975
DB 661 GAATTCAGGCGCGGTTTCAACAAGAGGCGAGGGACCTTGCCTACTATGAGATCTGTGACT 720

QY 976 tctccgcggagccacagtctcatalgaacacctcgccagcaggtccctctatgccaccaa-g 1034
DB 721 TCCTCCGGGAGGCCACAGTCCATAGAACTCTCGG-CAGCAGGTCCCTATGCCACCAAGG 779

QY 1035 ggcacacagtggtgagatcacgacgaccaggaaagcgctcaaaagcaaggtg 1085
DB 780 GGCAACAGTGGGTAGATACGACGACGACCAAGAAAGCGCTCAAGCAAGGTG 830

RESULT 6
BG741628 892 bp mRNA EST 15-MAY-2001
LOCUS 602635472F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4780673 5',
DEFINITION mRNA sequence.
ACCESSION BG741628
VERSION BG741628.1 GI:14052281
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 892)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA
Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10639 row: P column: 18
High quality sequence stop: 856.
Location/Qualifiers
1..892
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4780673"
/lab_host="DH10B (T1 phage-resistant)"
/Note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 212 a 248 c 236 g 196 t
ORIGIN
Query Match 45.4%; Score 762.8; DB 11; Length 892;
Best Local Similarity 94.1%; Pred. No. 1.9e-183;
Matches 839; Conservative 0; Mismatches 42; Indels 11; Gaps 4;
QY 9 ctggcacagagaccctggtggaagggagggagggccacacctgacctgctctgctgcagcc 68
DB 1 CTGGCACAGGACCCGTGGGCAAGGGAAGAGCCACACCCCTGCCCTGCTGCTGCGACGC 60
QY 69 agaattgggtgtgagggcgtctcaaacaggcttctgtggtcctgtgctgctcctcagtgctgc 128
DB 61 AGAATGGGTGTGAAGCGCTCTCAACAGGCTTGTGTGCTCTGCTGCTGCTGCTGCTGCTG 120
QY 129 tctgcataaacctgctgtctgtactacacagctggtccagtagcagggagggagggatggg 188
DB 121 TCTGCATACAAACTGTGTCTACTACACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 180
QY 189 agctgtctccagatgacctgacctgacctgacctgacctgacctgacctgacctgacctg 248
DB 181 AGTGTCTCCAGATGCTTGTGACCGCTTGTGACCGCTTGTGACCGCTTGTGACCGCTTGTG 240
QY 249 aataagcaacatcacatcgacacctggagtggaatgatgtgacgctctcagggcagtg 308
DB 241 AATAAAGCAACATCATCGACACCTGGGAGTGGGAATGATGTGACGCTCTACGGCATG 300
QY 309 ctcaacacactcaaacacagaccccccaacctgaag--actctctgtctgtcgaggat 366
DB 301 CTCAACACACTCAAGAACAGGAACCCCAACTTGAGGAATCCTCTGTGCTGGCGGAGGTA 360
QY 367 ggaactttgggt-----ctcaagaattttcccaagatagcctcccaacacacagagtcgc 420
DB 361 TGGACCTTGGGTGCTCTCAAGAGTATTTTCCAAAGTATAGCTTCCAAACCCAGAGTGC 420
QY 421 ggaatttcacagtgtagtaccgccatttctgcgacccatggtcttgaatgagggcgtagcc 480
DB 421 GGACTTTCAATCAAGTCAGTACCGCCATTTCTGCGACCCATGCTTTTATGATGGGTGGAC 480
QY 481 ttgcctggtctctaccctggagcgagagagacacacacattttaccacccctaatcaagaaa 540
DB 481 TTGCCTGGCTCTACCTTGGACGGAGAGACAAACAGACATTTTACACCCCTAATCAAGAAA 540
QY 541 tgaagggcgaattataaaggaagccagccaggggaaagagcagctcctctcagcgag 600
```

```

Db 541 TGAAGCCCAATTATTAAGGAGAGCCAGCCAGGGNAAAGCAGCTCTGCTCAGGCGAG 600
QY 601 cactgtctcggggaaggtcaacattgacagcagctatgacattgccaagatatcccaac 660
Db 601 CACTGTCTGGGGGAAGGTCACCATTTGACAGCAGCTATGACATTGCCAAGATATCCCAAC 660
QY 661 acctggatttcattagcattacagcattacagattttatcgccctgctgctggagaccag 720
Db 661 ACCTGGATTTCATTAGCATCATGACCTAGCATTTTTCATGAGCCTGGCGTGGGACACAG 720
QY 721 gcatcacagtcctccagcaggtccagaggtcagaggtacagagtcctgacagattcagcaaca 780
Db 721 GCATCACAGTCCCTCTGTCGAGGTCAGAGGATGCAAGTCTGACAGATTCAGCAACA 780
QY 781 ctgactatgctgggtgacatgtt--gaggctggggctcctccagtaagc-tggtga 837
Db 781 CTGACTATGCTGTGGGTACATCTTTGAGGCTGGGGCTCTCCAGCATTAAGCTTGTGA 840
QY 838 tgggcatcccaactctgggagagagcttcaactctgcttcttctgagactgg 889
Db 841 TGGGCATTCCCAACTTGGGAGGAGCTTCACTCTGGGTTCTTCTGAGACTGG 892

RESULT 7
BF338858 885 bp mRNA EST 22-NOV-2000
LOCUS 602036114F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4184187
DEFINITION 5', mRNA sequence.
ACCESSION BF338858
VERSION BF338858.1 GI:11285277
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 885)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: gcapbs-r@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LHAM9501 row: k column: 04
High quality sequence stop: 739.
Location/Qualifiers
1. 885
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4184187"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 204 a 251 c 239 g 190 t 1 others
ORIGIN

Query Match 44.8%; Score 752.4; DB 11; Length 885;
Best Local Similarity 94.9%; Pred. No. 8.3e-181;
Matches 799; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

QY 7 agctggcaccagagccgtggggaagggaagggcagaccctgctgctgctgcag 66
Db 1 AGCTGGCACCAGAGCCGTGGGAGGAGGAGGACACACCTGCTGCTGCTGCAG 60

```

```

QY 67 ccagaatgggtgtgaaggcgtctcaaacagggttttggctcgtgggtgctccagtgt 126
Db 61 CCAGAATGGGTGTGAAGGCGTCTCAAACAGGCTTTGTGTCTCTGCTGCTCCAGTGCT 120
QY 127 gctctgcatacaaaactggtctgctactacacacagctggtccacgtaccgggagggcgtg 186
Db 121 GCTCTGCATACAAACTGGTCTGCTACTACACAGCTGGTCCCAGTAGTACCGGAGGCGATG 180
QY 187 gtagctgtctccacagatgcccctgacgcttctggtgtaccacacatcatcactacagcttgg 246
Db 181 GGAGCTGCTTCCAGATGCCCTTGACCGCTTCTCTGTGTACCCACATCATCTACAGCTTGG 240
QY 247 ccaataaagcaacgacatcacatcgacacctgggagtggaatgatgtgacgtcttaaggca 306
Db 241 CCAATATAGCAACGATCATCATCGACACCTGGGAGTGGATGATGTGACGCTCTACGGCA 300
QY 307 tgcacaacacactcaacaacacacacacacacacacacacacacacacacacacacacacac 366
Db 301 TGCTAACACACTCAAGAACAGGAACCCCAACCTGAAGACTCTCTTGTCTGTCTGGAGGAT 360
QY 367 ggaacttgggtctcaaaagattttccaagatagcctccaacacccagagtcgcgcgactt 426
Db 361 GGAACCTTTGGGTCTCAAGATTTTCCAAGATAGCTTCCAACACCCAGAGTCGCCGGACTT 420
QY 427 tcatcaagtcagtcaccgcccatttctgcgaccccatggttggatggcgctgacactggcct 486
Db 421 TCATCAAGTCAGTACCCCAATTTCTGGCACCCCATGGCTTTGATGGGCTGACCTTGCCT 480
QY 487 ggtctaccctggagcgagagacaaacacacacacacacacacacacacacacacacacacac 546
Db 481 GGCTTACCTTGGAGCGGGAGACAAACAGCATTTTACCACCTTAATCAAGGAATGAAGG 540
QY 547 ccgaattataaaggagagccagccaggggaaaaagacgtcctcgtcagcgcacactgt 606
Db 541 CCGAATTTATAAGGAAGAGCCAGCCAGGAAAAGCAGCTCTGCTCAGCGCAGCATGT 600
QY 607 ctgctggggaaggtccaccatgacagcagctatgacattgccaagatatcccaacacactgg 666
Db 601 CTGCGGGGAAGGTCCACCTTGACAGCAGCTATGACATTTGCCAAGATATCCCAACACCTGG 660
QY 667 attcattagcatcatgacctacgattttcattgctggcgtggcggtgggacacacagccatc 726
Db 661 ATTTCAATAGCATCATGACCTACGATTTCATGAGCGCTGGCGTGGGACCA-AGGTGATC 719
QY 727 acagtcctccctcagcgaggtcagaggtgcaagtctcctcagcagattcagcacaactgact 786
Db 720 ACAGTCCCTGTTCCGAGGTCAGAGGATTTGCCAGTCTCTGACAGATTTTCGCAA-ACTGGCT 778
QY 787 atgctgtgggtacatgttgaggctggggggtcctcgtccagtaagctggtgtagggcatcc 846
Db 779 ATGCTGGGGGTCTCTTGTGTGACGCGGGGGTCTCTGCCATTAAAGCTGTGATGGGTCCC 838
QY 847 cc 848
Db 839 AC 840

RESULT 8
BG696335 907 bp mRNA EST 07-MAY-2001
LOCUS 602659494F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4802830 5',
DEFINITION mRNA sequence.
ACCESSION BG696335
VERSION BG696335.1 GI:13961375
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 907)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

```

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
Sequencing by: The I.M.A.G.E. Consortium (LLNL) DNA
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Place: LLMI0697 row: k column: 23
High quality sequence stop: 825.
Location/Qualifiers
1. .907
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:4802830"
/clone_lib="NCI_CGAP_Skn3"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
218 a 251 c 240 g 198 t

FEATURES
source

Query Match 44.4%; Score 746.8; DB 11; Length 907;
Best Local Similarity 95.2%; Pred. No. 2.2e-179;
Matches 847; Conservative 0; Mismatches 32; Indels 11; Gaps 7;

QY 2 taggtagctggcaccagagcgtggcgaaggaaagccacacacctgctctgc 61
|||||
DB 1 TAGGTAGCTGGCACAGGAGCGGTGGGCAAGGAAGGCCACACCTGCCCTGCTGC 60
|||||

QY 62 tcagcgcagaatgggtgtgaaggcgtctcaaacagcgtttgtgctcgtgctcca 121
|||||

DB 61 TGCAGCAGCAATGGGTGTGAAGCGCTCTCAACAGCGCTTTGTGGCTGGTGCTCCA 120
|||||

QY 122 gtgctcgtcgtacaaactggctgtctactacacacagctggctccagctacgggaag 181
|||||

DB 121 GTGCTGCTGTGCATACAAACTGTGCTGTACTACACAGCAGTGTGTCCAGTACCGGAAG 180
|||||

QY 182 cgaatggagctgctccagatgctctgaacgctcctgtgctgtacacacatcatcac 241
|||||

DB 181 CGATGGAGCTGCTCCAGATGCCCTTGACCGCTTCTCTGTACCCACATCATCTACAG 240
|||||

QY 242 ctttgccaataagaacagatcacatcgacacctgggagtggaatgatgtgacgctcta 301
|||||

DB 241 CTTTGCCAATATAAGCAACGATCACATCGACACCTGGGAGTGGAAATGATGTGACGCTCTA 300
|||||

QY 302 cggcgtcgtcaacacactcaaacacagcaaccccaacctgaagactctctgtctcgg 361
|||||

DB 301 CGGCATGTCTCAACACACTCAAGAACAAGAACCCCAACCTGAAGACTCTCTGTGCTCGG 360
|||||

QY 362 agatggaacttggctcgaagatcttccaagatgctcctcaacacacagagtcgag 421
|||||

DB 361 AGATGGAACCTTTGGGTCTCAAGATTTTCCAGATAGCTTCCACACCCAGAGTGC 420
|||||

QY 422 gacttcacagtcagtcgcgcattcttgcgcacccatggcttggatgggctgacct 481
|||||

DB 421 GACTTTTCATCAAGTCAGTACCGCCCAATTTCTGCGCACCCCATGGCTTTGATGGGTGGACCT 480
|||||

QY 482 tgctggctctacctggagcgagagacaacacacattttaccacacctaatcaagaaat 541
|||||

DB 481 TGCTGTGCTGTACCTCTGGAGGAGAGACAAACAGCATTTTACCACCTAATCAAGAAAT 540
|||||

QY 542 gaagccgaatttataaaggaaagccagcaggaggaagaaagcagctcgtcctcagcagc 601
|||||

DB 541 GAAGGCCGAATTTATAAGGAAGAACCCAGCAGGAAAGAGACAGCTCTCTGCTACGCGCAGC 600
|||||

QY 602 actgctgcgggggaaggttcaccattgacagcagctatgacattgccaagatatcccaaca 661
|||||

DB 601 ACTGTCTCGGGGAAGTCAACCATTTGACAGCAGCTATGACATTGCCAAGATATCCCAACA 660
QY 662 cctggatttcattagcatcatgacacctacattttcattgagcgcctggcgtggga--ccaca 719
|||||
DB 661 CTTGGATTTCATTAGCATCATGACCTACCATTTTCATGGAGCCTGGCTGGGAACACACAA 720
|||||
QY 720 ggcacatcagtcctccctcag--gcgaggtcagg--aggatgcaagctcctga--cagatttc 773
|||||
DB 721 GGCATCAACAAGTCCCTCCCTGTTCCCGAGGTGTCAGGGAGGATGCAAGTCTCTGACCATTC 780
|||||
QY 774 agcaacact-gactatgc-tgtgggggtaca-tgttgaggtcgggggctcctcgtccagtaag 830
|||||
DB 781 AGCAACACTGGACTATGCTTGTGGGTACATTTGTAGGCTGGGGCTCTGACAGGTAAG 840
|||||

QY 831 ctggtgatgggcatcccaaccttcgggagagcttcacactcgtgctcttc 880
|||||
DB 841 TGTTGTATGGGCAATCCACCTTCGGGAGGAGCTTCTCTTGGCTTCTTC 890
|||||

RESULT 9

LOCUS BG740681 882 bp mRNA EST 15-MAY-2001
DEFINITION 602631527F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:4776582 5',
mRNA sequence.

ACCESSION BG740681.1 GI:14051334

VERSION BG740681

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 882)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: James Cleaver, M.D.

cDNA Library Preparation: Life Technologies, Inc.

Sequencing by: The I.M.A.G.E. Consortium (LLNL) DNA

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLMI0629 row: f column: 07

High quality sequence stop: 780.

FEATURES

source

1. .882

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone_image="IMAGE:4776582"

/clone_lib="NCI_CGAP_Skn3"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: skin; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

BASE COUNT 219 a 243 c 243 g 177 t

ORIGIN

Query Match 44.4%; Score 746.2; DB 11; Length 882;
Best Local Similarity 96.1%; Pred. No. 3.1e-179;
Matches 787; Conservative 0; Mismatches 28; Indels 4; Gaps 2;

QY 253 taagcaacatcacatcgacacctgggagtggaatgatgtgacgctctacgcatgctca 312
|||||

DB 1 TAAGCAACGATCACATCGACACCTGGAGTGGAAATGATGTGACGCTCTACGGCATGCTCA 60
|||||

QY 313 acacactcaacaacacgaaccccccaacctgaagactctcttctgtcgtggaggatggaact 372
|||||

DB 61 ACACACTCAAGAACAGGAACCCCAACCTGAAGACTCTCTTGTGTGTCGGAGGATGGAAC 120
|||||

| | | | |
|------------|-------------|----------------------------------------------------------------|---------------------------------------------------------------|
| QY | 373 | ttgggtctcaaaagattttccaaagatagcctccaaacccagagtcgcccgaactttcatca | 432 |
| Db | 121 | TTGGGTCTCAAGATTTCCTCAAGATAGCTTCCAAACCCAGAGTCGCCGACTTTCATCA | 180 |
| QY | 433 | atcagtcacgcacatttctgcacccatgcttgaatggcgtgacctggcctgctct | 492 |
| Db | 181 | AGTACGTACCGCATTTCTCGGCACCCATGCGCTTTGATGGCTGGACCTGCTGGCTCT | 240 |
| QY | 493 | accttgacgagagagacaaacacccattttaccacccctaatcaaggaaatgaagccgaat | 552 |
| Db | 241 | ACCTGGACGGAGAGACAAACAGCATTTTACCACCTTAATCAAGGAAATGAAGCCGAAT | 300 |
| QY | 553 | ttataaaggaagccagccaggggaaagacagctctctgctgcagcagcaactgctgcgg | 612 |
| Db | 301 | TTATAAAGGAAGCCAGCCAGGAAAGAGCAGCTCTCTCAGCGCAGCACTGCTCTCGG | 360 |
| QY | 613 | ggaaggtcacattgacagcagctatgacattgccaagatattcccaacactgatttca | 672 |
| Db | 361 | GGAGGTTCACCATTTGACAGCAGCTATGACATTCGAAGATATCCCAACACTGGATTTC | 420 |
| QY | 673 | ttagcatcatgacacagatttttctgctgctggcgtgggacacagggccatcacagtc | 732 |
| Db | 421 | TTAGCATCATGACCTACGATTTTCATCGAGCCTGGCGTGGACACAGGCCATCACAGTC | 480 |
| QY | 733 | ccctcagcgcaggtcagggatgcaagtcctgacagattcagcaacactgactatgctg | 792 |
| Db | 481 | CCCTGTTCGAGGTCAGGAGGATCAAGTCCTGACAGATTTCAGCAACACTGACTATGCTG | 540 |
| QY | 793 | tgggtacatgtgagctggggctcctgccagtaagctgggtgagcgcacccacact | 852 |
| Db | 541 | TGGGGTACATGTTGAGGCTGGGGGCTCCTGCCAGTAAGCTGGTGGATGGGCATCCCACT | 600 |
| QY | 853 | tcgggagagcttcaactctgctctctgagactggtttccagcggccaatctccaggac | 912 |
| Db | 601 | TCGGGAGAGGCTTCACCTGCTTCTTGAGACTGGTGTGGAGCCCAATCTCAGGAC | 660 |
| QY | 913 | cggaattccagggccggttccaaagagggcagggaccccttgctactatgactatgctg | 972 |
| Db | 661 | CGGGAATTCAGGCGGTTTACCAAGGAGGCGAGGAGCCCTTGCCCTACTATGAGATCTGTG | 720 |
| QY | 973 | acttccctcggagcagacagtcacagaa--ccctcggccagcaggtccccc--tatgcc | 1028 |
| Db | 721 | ACTTCCCTCCGCGGAGCCACAGTCCATAGATCTCTTGGCCAGCAGAGTCCCTCTATGCCAC | 780 |
| QY | 1029 | accgaaggcaaccagtggtgaggtacagacaccagga | 1067 |
| Db | 781 | CCAAAGGCAACACAGGTGGTAGGATACACACACACAGGA | 819 |
| RESULT 10 | | | |
| BG697544 | | | |
| LOCUS | 602660764F1 | NCI_CGAP_Skn3 | Homo sapiens cDNA clone IMAGE:4804071 5', |
| DEFINITION | 602660764F1 | NCI_CGAP_Skn3 | Homo sapiens cDNA clone IMAGE:4804071 5', |
| ACCESSION | BG697544 | | EST |
| VERSION | BG697544.1 | | GI:13963875 |
| KEYWORDS | | | mRNA sequence. |
| SOURCE | | | human. |
| ORGANISM | | | Homo sapiens |
| REFERENCE | | | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; |
| AUTHORS | | | Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. |
| TITLE | | | 1 (bases 1 to 976) |
| JOURNAL | | | NIH-MGC http://mgc.nci.nih.gov/. |
| COMMENT | | | cDNA Library Prepared by: The I.M.A.G.E. Consortium (MGC) |
| | | | Unpublished (1999) |
| | | | Contact: Robert Strausberg, Ph.D. |
| | | | Email: cgapbs-re@mail.nih.gov |
| | | | Tissue Procurement: James Cleaver, M.D. |
| | | | cDNA Library Arrayed by: The I.M.A.G.E. Consortium |
| | | | Sequencing by: Incyte Genomics, Inc. |
| | | | Clone distribution: MGC clone distribution information can be |
| | | | found through the I.M.A.G.E. Consortium/LLNL at: |

| | | | |
|----------------------------------------------------------------|------------------------------------------------------------------------------------------------------------------|-----------------------------------------------------------------|------|
| http://image.llnl.gov | | | |
| Plate: LLAM10700 row: o column: 16 | | | |
| High quality sequence stop: 837. | | | |
| Location/Qualifiers | | | |
| FEATURES | source | | |
| | 1. .976 | | |
| | /organism="Homo sapiens" | | |
| | /db_xref="taxon:9606" | | |
| | /clone="IMAGE:4804071" | | |
| | /lab_host="NCI_CGAP_Skn3" | | |
| | /note="Organ: skin; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dm. | | |
| | Average insert size 1.5kb. Library constructed by Life | | |
| | Technologies. Note: this is a NCI_CGAP Library." | | |
| BASE COUNT 217 a 289 c 273 g 197 t | | | |
| ORIGIN | | | |
| Query Match 43.7%; Score 733.8; DB 11; Length 976; | | | |
| Best Local Similarity 91.2%; Pred. No. 4.6e-176; | | | |
| Matches 880; Conservative 0; Mismatches 72; Indels 13; Gaps 9; | | | |
| QY | 352 | tgctgtcggagatggaacttttggtctcaaaagattttccaaagatagcctccaacaccc | 411 |
| Db | 2 | TGCTGTGCGAGGATGGAACCTTGGGTCTCAAGATTTTCCAGATAGCCTCCCAACACC | 61 |
| QY | 412 | agagtcgcccgaacttca--tcaagtcagtaacccgacattctgcgacccatggtttgat | 470 |
| Db | 62 | AGAGTCGCGGACTTTTCAGTCAAGTCAGTACCGGCATTTCTGCGACCCATGGCTTTGAT | 121 |
| QY | 471 | ggcggtgacctgctgctgctacccctggagagagacacacacacacacacacacaccta | 530 |
| Db | 122 | GGGCTGACCTTGCTGGCTTACCCCTGGAGGAGACAAACAGCATTTTACCACCCCTA | 181 |
| QY | 531 | atcaagaaatgaaggccgaatttataaagaaagccagccaggggaaaaaagcagctctg | 590 |
| Db | 182 | ATCAAGGAATGAGGCGGAATTTATAAGGAAGCCAGCCAGGAGGAGGAGGAGGAGG | 241 |
| QY | 591 | ctcagcgcagcactgtctggggaaaggtaccattgacagcagctatgacatgcgaag | 650 |
| Db | 242 | CTCAGCAGCAGCAGTCTGCGGGGAGGTCACCATTTGACAGCAGCTATGACATTTGCCAAG | 301 |
| QY | 651 | atattcccaacacctggatttca--ttagcatcatgacctacattttcatggcctggcg | 709 |
| Db | 302 | ATATCCCAACACCTTGGATTTCCTTAGCATCATGACCTACGATTTTCATGAGCCTGGCG | 361 |
| QY | 710 | tgggaccacagggccatcacagtcctcctcagcaggtgcaggtgcaggtgcaggtgcaggt | 769 |
| Db | 362 | TGGGACCACAGGCGCATCACAGTCCCTTGTCCGAGGTGAGGAGGATGCAAGTCCCTGACAG | 421 |
| QY | 770 | attcagaacacactgac--tatgtgtgggtgtacatgtgtgagctgggggctcctccagta | 828 |
| Db | 422 | ATTGAGCAACACTGACGTATGCTGTGGGGTACATGTTGAGGCTGGGGGCTCCTGCCAGTA | 481 |
| QY | 829 | agctgggtgagtgatcctccacacctcgggagaggttcaactctggtcttctctgagactg | 888 |
| Db | 482 | AGCTGGGTGATGGGCATCCCCACCTTCGGGAGGAGGCTTCACCTCTGCTTCTTCTGAGACTG | 541 |
| QY | 889 | gtgttccagcggcccaatctcagggacgggaattccaggcgggttccaccaagggagcgagga | 948 |
| Db | 542 | GTGTGGAGGCCCCAATCTCAGGACCGGGAATTCAGGCGCGGTTCACCAAGAGGAGCGGGA | 601 |
| QY | 949 | cccttgctactatgagatctgtgacttctcccgaggagccac--agtcacatagaacacct | 1007 |
| Db | 602 | CCCTTGCTTACTATGAGATCTGTGACTTCTCCCGGAGGAGCCACAAAGTCCATAGAAATCTC | 661 |
| QY | 1008 | ggcagcaggtccctctatgccaccaagggcaacagtggtggtaggatacagacagcagga | 1067 |
| Db | 662 | GGCCAGCAGGTCCCTTATGCCACCAAGGGCAACAGTGGGTAGGATACGACAGCCAGGA | 721 |
| QY | 1068 | agcgtcaaaagcaaggtgcagta--cctgaagataggtgcaggtgcaggtgcaggtgcaggt | 1124 |
| Db | 722 | AGCGTCAAAAGCAAGGTGAGTGTGCTGAAGGACAGGAGGAGGAGGAGGAGGAGGAGGAGG | 781 |

| | Query Match | 42.3%; | Score 711.4; | DB 11; | Length 912; |
|----|-----------------------|------------------------------------------------------------------|---------------------|----------------|-------------|
| | Best Local Similarity | 91.4%; | Prod. No. 2.3e-170; | Indels | 9; Gaps |
| | Matches 800; | Conservative | 0; | Mismatches 66; | |
| QY | 12 | gcaccagagcgtgtggcaaggaagagccacacccctgcctgctctgtctgcagccaga | 71 | | |
| DB | 1 | GCACCAGAGCGGTGGGAAGGAAGAGCCACACCCCTGCCCTGCTCTGCTGCACCCAGA | 60 | | |
| QY | 72 | atgggtgtgaagcgctctcaaacagcgcttttggctcctgggtgctctccagtgctgctct | 131 | | |
| DB | 61 | ATGGGTGTGAAGCGCTCTCAACAGCGCTTTGTGGTCTTGCTGCTCCAGTGCCTGCTCT | 120 | | |
| QY | 132 | gcatacaaaactggtctgtactacaccagctggtcccaagtaccgggaagggcagatgggagc | 191 | | |
| DB | 121 | GCATACAAACTGGTCTGTCTACTACACCAGCTGGTCCCACTACCGGGAAGGCCATGGGAGC | 180 | | |
| QY | 192 | tgtctccagatgccctgacgccttcctgtgtaccacacatcatctacagctttgcgaat | 251 | | |
| DB | 181 | TGCTTCCCAAGATGCCCTTGACCGCTTCCTCTGTACCCACATCATCTACAGCTTTGCCAAT | 240 | | |
| QY | 252 | ataagcaacgatacacatcgacacacctgggagtgatgtgacgctctacggcatgctc | 311 | | |
| DB | 241 | ATAAGCAACGATCACATCAGACACCTGGGAGTGGAAATGTTGACGGCTCTACGGCATGCTC | 300 | | |
| QY | 312 | aacacactcaacaacacgacacccccacacctgaagactctcttctgtctggagagatggaaac | 371 | | |


```
Db 301 AACACACTCAAGAACAGGAAACCCCAACCTGGAAGACTCTCTGTCTCTCGAGGATGGAAC 360
Qy 372 ttgggtctcaagaattttccaaagatagctctcaacacccagagtcgcgcgagcttctc 431
Db 361 TTTGGGTCTCAAGATTTTCCAAAGATAGCTCTCAACACCCAGAGTCGCCGAGATTTCATC 420
Qy 432 aagtcagtagccaccattctgcgcacccatggtctttgatggcgtagaccttgcctggctc 491
Db 421 AAGTCAGTACCGCCATTCTGCGCACCCATGCTTTGATGGGTGGACCTTGCTGGCTC 480
Qy 492 taccctggcggagagagacacacacattttaccacccctaatcaaggaaatgaaggccgaa 551
Db 481 TACCCTGGACGGGGAGACAAACAGCATTTTACCACCTTAATCAAGGAAATGAAGGCCGAA 540
Qy 552 titataaaggagccagccaggggaaagagcagctctctcagcagcagcactgtctgcg 611
Db 541 TTTATNAAGAGAGCCAGCCAGCGGAAAGAGAGCTCTCTCTCAGCCGACGACTGTCTGG 600
Qy 612 gggaaagtcaccattgacagcagctatgacattgccaagatatcccaacacacctggatttc 671
Db 601 GGAAGGTCAACATTGACAGCAGCTATGACATTTGCCAAGATATCCCAAAACCTGGATTTC 660
Qy 672 attagcatcatgacctacgattttcatg---gogcctggcgctgggaccacagccatcac 728
Db 661 ATTAGCATCATGACCTTACGATTTTCTATGGAGCTCGCGCTGGGGACACAGGGGCCATCAC 720
Qy 729 agtcccc---tcaggcaggctcagagatgcaagtctctgacag-attcagcaacactg 783
Db 721 AGTTCCTCCCTGTTCCAGAGTTCAGGAGCAGCTGCTCTCTGATGAGAAATCACAGCATG 780
Qy 784 actatgctggggtacatgttgaggctggggctcctccagtaagctggtgatggcga 843
Db 781 GCTATGCGGTGGGGTCTCTGTTGCGCGGGGGGGCGCGGACGTAACCTGGTATGGGCA 840
Qy 844 tccccacctcgggagagcttccactctggtctct 878
Db 841 T-CCCACCTTCGGGGGGGCTTCACTCGGCTCTCT 874

RESULT 14
BF339902
LOCUS
DEFINITION BF339902 819 bp mRNA EST 22-NOV-2000
5', mRNA sequence.
ACCESSION BF339902
VERSION BF339902.1 GI:11286363
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 819)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-re@mail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9502 row: 9 column: 05
High quality sequence stop: 717.
Location/Qualifiers
1..819
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:418476"
/clone_lib="NCI_CGAP_Brn64"
```

FEATURES
SOURCE

RESULT 15
BF339615

```
BASE COUNT 203 a 230 c 215 g 171 t
ORIGIN
Query Match 42.0%; Score 706; DB 11; Length 819;
Best Local Similarity 95.5%; Pred. No. 5.2e-169;
Matches 781; Conservative 0; Mismatches 30; Indels 7; Gaps 5;

Qy 13 caccagagccgtgggcaaggagggccacacccctgccctgctctgctgcagccagaa 72
Db 1 CACCAGAGCCGTGGCAGAGGAGGAGGCCACACCTCGCTGCTGCTGCGAGCAGAA 60
Qy 73 ttgggtgtaaggcgtctcaaacaggctttgtgctcctggctgctgctccagtgctctg 132
Db 61 TGGGTGTGAAGGCGTCTCAAAACAGGCTTAGTGTCTCTGGTGTCTCCAGTGTGCTCTG 120
Qy 133 catacaaacctgctctactacacacagctggtccacagtagccgggagggagctgagct 192
Db 121 CATACAAACTGGTCTGTCTACTACACAGCTGTGTCCAGTACCAGGAGGCGATGGGAGCT 180
Qy 193 gcttccagatgccctgaccgcttctctgtgtaccacacatctacagcttgcceata 252
Db 181 GCTTCCAGATGCCCTTGACCGCTTCTCTGTACCCACATCATCTACAGCTTTGCCAATA 240
Qy 253 taagcaacgatcatcagacacccctgggagtggaatgatgtgacgctctacggcatgctca 312
Db 241 TAAGCAACGATCATCGACACCTGGAGGTGGAATGATGTGACGCTCTACGGCATGCTCA 300
Qy 313 acacatcaacaacacgaaaccccaacctgaagactctctgtctgtcgaggagatgaaact 372
Db 301 ACACACTCAAGAAACAGAGAACCCCAAGCTGAAGACTCTCTTGTCTGTCGGAGATGGAAT 360
Qy 373 ttgggtctcaagattttccaagatagctctcaacacccagagtcgcgcgacattca 432
Db 361 TTGGGTCTCAAGATTTTCCAAAGATAGCTTCCAAACACCCAGAGTCCGCGGACTTTATCA 420
Qy 433 agtcagtaccgcctattctgcgcacccatggtttgatggcgctgacctgacctgctct 492
Db 421 AGTCAGTACCGCATTTCTGCGCACCCATGGCTTTGATGGCTGGACCTTGCCTGGCTCT 480
Qy 493 accctggcggagagagacaaacacacattttaccaccttaatacaaggaaatgaaggccgaat 552
Db 481 ACCCTGGAGGGGAGACAAACAGCATTTTACCACCTTAATCAAGGAAATGAAGGCCGAAT 540
Qy 553 ttataaaggagccagccaggggaaaaagcagctcctgctcagcgcagcactgtctcgg 612
Db 541 TTATAAAGGAAGCCCGCCAGCCAGGAAAGAAAGAGCAGCTCTCTGCTCAGCGGAGCACTGTCTGCGG 600
Qy 613 ggaaggtcaccattgacagcagctatgacattgccaaagatattcccaacctggatttca 672
Db 601 GGAAGTCAACATTGACAGCAGCTATGACATTGCCAAGATATCCCAACACCTGGATTCA 660
Qy 673 ttatcatcatgacctacgattttcatg---cgctctggcgtgggaccacagcccatcac-- 728
Db 661 TTAGCATCATGAGTACGATTTTTCATGGGAGCCTGGCGTGTGGGACCAAGCCATCACCA 720
Qy 729 agtccccctcagggagtcagagagatgcaagctct-gacagattcagcagcact-gact 786
Db 721 GTCCCCCATGTTCCAGAGTCAAGGAGGATGCAAGTCTCTGGAAGATCCAGCAACACTGGACT 780
Qy 787 atgctgtgg--ggtacatgttgaggctgggggctcctctgc 823
Db 781 ATGCTGGGGCGGTACATGTTGAGGCTGGGGGATCTGGC 818
```

/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

LOCUS BF339615 906 bp mRNA EST 22-NOV-2000
DEFINITION 602039090F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4186727
5', mRNA sequence.
ACCESSION BF339615
VERSION BF339615.1 GI:11286070
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 906)
REFERENCE NIH-MGC http://mgs.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabp@remail.nih.gov
Tissue procurement: David N. Louis, M.D.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9508 row: d column: 24
High quality sequence stop: 730.
Location/Qualifiers
1. 906
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4186727"
/clone_lib="NCI_CGAP_Brn64"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (11 phage-resistant)"
/note="Organ: Brain; Vector: PCWV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 200 a 264 c 255 g 187 t
ORIGIN
Query Match 41.2%; Score 692.8; DB 11; Length 906;
Best Local Similarity 95.5%; Pred. No. 1.2e-165;
Matches 745; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
QY 155 caccagctggtccagtcaccgggaaggcgatggagctgtccagatgcccttgaccg 214
|||||
Db 1 CACCAGCTGGTCCCACTACCGGGAAGGCGATGGGAGCTGTCTCCAGATGCCCTTGACCG 60
QY 215 ctctctgtaccacatcatctacagcttttgcataataagaacacgacatcacatgcacac 274
|||||
Db 61 CTCTCTGTATCCCACTATCATCTAGCTTTCCTAATAAGCAACGATCACATCGACAC 120
QY 275 ctgggagtggaatgtagctgacgcttaagcattgctcaacacacatcaacacgaaccc 334
|||||
Db 121 CTGGGAGTGAATGATGACGCTCTACGGCATGTCTACACACATCAAGAACAGGAACCC 180
QY 335 caacctgaagactctcttctgtcgagagatggaacttgggtctcaagaattttccaa 394
|||||
Db 181 CAACCTGAAGACTCTC-TGTCTGTCGGAGGATGGAAC-TTGGGTCTCAAGATTTTCCAA 238
QY 395 gatagctccaacacacagagtcgcggagctttcatcaagtcagtcaccgcatcttctg 454
|||||
Db 239 GATAGCCTCCAACACCCAGAGTCGCCGGACTTTCATCAAGTCAGTACCGCCATTTCTGG 298
QY 455 caccatggctttgatggcggtgaccttgcctggctctaccctggcgagagacaaaca 514
|||||
Db 299 CACCATGGC-TTGATGGCTGGACCTTGCTGGCTCTACCCCTGGACGGAGAGACAAACA 357
QY 515 ccattttaccacctaatcaagaatgaagcgccgaatttatataaaggagcccgagcagg 574
|||||
Db 358 GCATTTTACCACCTAATCAAGGAATGAAGGCCGGAATTTATNAAGGAAGCCCGCAGG 417

QY 575 gaaaaagcagctctgctcagcgacgacgtgtctgcggggaagtgaccattgacagcag 634
|||||
Db 418 GAAAAAGCAGCTCCTGCTCTCAGCGGAGCAGCTGTCTGCGGGGAGGTCACCATTTGACAGCAG 477
|||||
QY 635 ctatgacatigccaagatatcccaacacacctggatttattagcatcatgacctacgattt 694
|||||
Db 478 CTATGACATTTGCCAAGATATCCCAACACCTGGATTTCATTAGCATCATGACCTACGATT 537
|||||
QY 695 tcatggcgcttggtggggaccagcccatcacagtcacagtcacctcagcgaggtcaggagga 754
|||||
Db 538 TCATGGAGCCTTGGGCTGGGACACAGGCCATCACAGTCCCTCTGTCGAGGTGAGGAGGA 597
|||||
QY 755 tgcgaagtctgacagattccagacactgactatgctgtgggtgtacatgttgaggtcgtg 814
|||||
Db 598 TGCAAGTCCTGACAGATTTCAGCAACACTGACTATGCTGTGGGTACATGTTGAGGCTGGG 657
|||||
QY 815 ggctctgcccagtaagctgggtgatgggcatccccaccttcgggaggaagcttcactctgc 874
|||||
Db 658 GGCTCCTGCCAGTAAGCTGGTGTGATGGGCATCCCACTTCGGGAGGAGCTTCACCTCTCGG 717
|||||
QY 875 ttctctgagactggtgttccagcgcccaatctcaggaacgggaattccaggccggttcac 934
|||||
Db 718 TTCTTCTGAGACTGGTGTGGGGGCCCAATCTCCGGAACGGGATTCCCGGCGCGGTCCC 777
|||||

Search completed: December 6, 2001, 22:52:45
Job time: 9225 sec

